P1 1131703

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

February 20, 2004

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM
THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK
OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT
APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A
FILING DATE.

APPLICATION NUMBER: *60/431,321*
FILING DATE: *December 06, 2002*
RELATED PCT APPLICATION NUMBER: *PCT/US03/38855*

```
RECEIVED
26 FEB 2004
WIPO      PCT
```
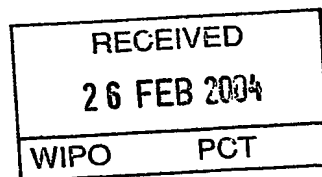
By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

P. SWAIN
Certifying Officer

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

BEST AVAILABLE COPY
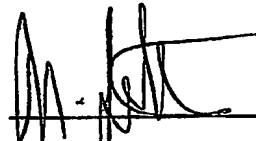
DEX – 0367

CERTIFICATE OF EXPRESS MAILING

"Express Mail" Label No. ET 657 238 006 US

Date of Deposit: <u>6 December 2002</u>

I hereby certify that this paper is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R 1.10 on the date indicated above and is addressed to the "BOX PROVISIONAL", Assistant Commissioner for Patents, Washington, D.C. 20231.

1) Provision Application Transmittal Letter, with authorization to charge deposit account 50-0434 $160.00 for filing fee (2 copies);

2) Application consisting of 232 pages including; 226 pages of Specification, 5 pages of Claims, and 1 page of Abstract; and

3) Return Post Card.

NATHAN P. LETTS
Registration No. 36,581

jc929 U.S. PTO
60/431321

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53 (c).

Express Mail Label No. ET 657 238 006 US

INVENTOR(S)

Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Roberto	Macina	San Jose, CA
Leah R.	Turner	Sunnyvale, CA
Yongming	Sun	Redwood City, CA

☐ Additional inventors are being named on the _____ separately numbered sheets attached hereto.

TITLE OF THE INVENTION (280 characters max)

COMPOSITIONS AND METHODS RELATING TO OVARIAN SPECFIFIC GENES AND PROTEINS

CORRESPONDENCE ADDRESS

Direct all correspondence to:

☐ Customer Number

OR   Type Customer Number here

Place Customer Number
Bar Code Label here

☒ Firm or Individual Name	Licata & Tyrrell				
Address	66 East Main Street				
Address					
City	Marlton	State	New Jersey	ZIP	08053
Country	USA	Telephone	(805) 810-1515	Fax	(805) 810-1454

ENCLOSED APPLICATION PARTS (check all that apply)

☒ Specification Total Number of Pages 232   ☐ CD(s), Number

☐ Drawing(s) Number of Sheets   ☐ Other (specify)

☐ Application Data Sheet. See 37 CFR 1.76

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

☐ Applicant claims small entity status. See 37 CFR 1.27.

☐ A check or money order is enclosed to cover the filing fees

☒ The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 50-0434

☐ Payment by credit card. Form PTO-2038 is attached.

FILING FEE AMOUNT ($) 160.00

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒ No.
☐ Yes, the name of the U.S. Government agency and the Government contract number are: _____

Respectfully submitted,
SIGNATURE _____

TYPED or PRINTED NAME   Nathan P. Letts

TELEPHONE   (650) 246-6400

Date: 6 December 2002

REGISTRATION NO. (if appropriate) 36,581

Docket Number: DEX-0367

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

COMPOSITIONS AND METHODS
RELATING TO OVARIAN SPECFIFIC GENES AND PROTEINS

FIELD OF THE INVENTION

5   The present invention relates to newly identified nucleic acids and polypeptides
present in normal and neoplastic ovarian cells, including fragments, variants and
derivatives of the nucleic acids and polypeptides. The present invention also relates to
antibodies to the polypeptides of the invention, as well as agonists and antagonists of the
polypeptides of the invention. The invention also relates to compositions comprising the

10   nucleic acids, polypeptides, antibodies, post translational modifications (PTMs), variants,
derivatives, agonists and antagonists of the invention and methods for the use of these
compositions. These uses include identifying, diagnosing, monitoring, staging, imaging
and treating ovarian cancer and non-cancerous disease states in ovarian, identifying
ovarian tissue and monitoring and identifying and/or designing agonists and antagonists of

15   polypeptides of the invention. The uses also include gene therapy, therapeutic molecules
including but limited to antibodies or antisense molecules, production of transgenic
animals and cells, and production of engineered ovarian tissue for treatment and research.

BACKGROUND OF THE INVENTION

Cancer of the ovaries is the fourth-most cause of cancer death in women in the

20   United States, with more than 23,000 new cases and roughly 14,000 deaths predicted for
the year 2001. Shridhar, V. et al., *Cancer Res.* 61(15): 5895-904 (2001); Memarzadeh, S.
& Berek, J. S., *J. Reprod. Med.* 46(7): 621-29 (2001). The incidence of ovarian cancer is
of serious concern worldwide, with an estimated 191,000 new cases predicted annually.
Runnebaum, I. B. & Stickeler, E., *J. Cancer Res. Clin. Oncol.* 127(2): 73-79 (2001).

25   Unfortunately, women with ovarian cancer are typically asymptomatic until the disease
has metastasized. Because effective screening for ovarian cancer is not available, roughly
70% of women diagnosed have an advanced stage of the cancer with a five-year survival
rate of ~25-30%. Memarzadeh, S. & Berek, J. S., *supra*; Nunns, D. et al., *Obstet. Gynecol.
Surv.* 55(12): 746-51. Conversely, women diagnosed with early stage ovarian cancer

30   enjoy considerably higher survival rates. Werness, B. A. & Eltabbakh, G. H., *Int'l. J.
Gynecol. Pathol.* 20(1): 48-63 (2001). Although our understanding of the etiology of
ovarian cancer is incomplete, the results of extensive research in this area point to a
combination of age, genetics, reproductive, and dietary/environmental factors. Age is a

key risk factor in the development of ovarian cancer: while the risk for developing

ovarian cancer before the age of 30 is slim, the incidence of ovarian cancer rises linearly

between ages 30 to 50, increasing at a slower rate thereafter, with the highest incidence

being among septagenarian women. Jeanne M. Schilder et al., *Heriditary Ovarian Cancer:*

5   *Clinical Syndromes and Management*, in <u>Ovarian Cancer</u> 182 (Stephen C. Rubin &

Gregory P. Sutton eds., 2d ed. 2001).

   With respect to genetic factors, a family history of ovarian cancer is the most

significant risk factor in the development of the disease, with that risk depending on the

number of affected family members, the degree of their relationship to the woman, and

10   which particular first degree relatives are affected by the disease. *Id.* Mutations in several

genes have been associated with ovarian cancer, including BRCA1 and BRCA2, both of

which play a key role in the development of breast cancer, as well as hMSH2 and hMLH1,

both of which are associated with heriditary non-polyposis colon cancer. Katherine Y.

Look, *Epidemiology, Etiology, and Screening of Ovarian Cancer*, in <u>Ovarian Cancer</u> 169,

15   171-73 (Stephen C. Rubin & Gregory P. Sutton eds., 2d ed. 2001). BRCA1, located on

chromosome 17, and BRCA2, located on chromosome 13, are tumor supressor genes

implicated in DNA repair; mutations in these genes are linked to roughly 10% of ovarian

cancers. *Id.* at 171-72; Schilder et al., *supra* at 185-86. hMSH2 and hMLH1 are

associated with DNA mismatch repair, and are located on chromsomes 2 and 3,

20   respectively; it has been reported that roughly 3% of heriditary ovarian carcinomas are due

to mutations in these genes. Look, *supra* at 173; Schilder et al., *supra* at 184, 188-89.

   Reproductive factors have also been associated with an increased or reduced risk

of ovarian cancer. Late menopause, nulliparity, and early age at menarche have all been

linked with an elevated risk of ovarian cancer. Schilder et al., *supra* at 182. One theory

25   hypothesizes that these factors increase the number of ovulatory cycles over the course of

a woman's life, leading to "incessant ovulation," which is thought to be the primary cause

of mutations to the ovarian epithelium. *Id.*; Laura J. Havrilesky & Andrew Berchuck,

Molecular Alterations in Sporadic Ovarian Cancer, in <u>Ovarian Cancer</u> 25 (Stephen C.

Rubin & Gregory P. Sutton eds., 2d ed. 2001). The mutations may be explained by the

30   fact that ovulation results in the destruction and repair of that epithelium, necessitating

increased cell division, thereby increasing the possibility that an undetected mutation will

occur. *Id.* Support for this theory may be found in the fact pregnancy, lactation, and the

use of oral contraceptives, all of which suppress ovulation, confer a protective effect with respect to developing ovarian cancer. *Id.*

Among dietary/environmental factors, there would appear to be an association between high intake of animal fat or red meat and ovarian cancer, while the antioxidant

5   Vitamin A, which prevents free radical formation and also assists in maintaining normal cellular differentiation, may offer a protective effect. Look, *supra* at 169. Reports have also associated asbestos and hydrous magnesium trisilicate (talc), the latter of which may be present in diaphragms and sanitary napkins. *Id.* at 169-70.

Current screening procedures for ovarian cancer, while of some utility, are quite

10   limited in their diagnostic ability, a problem that is particularly acute at early stages of cancer progression when the disease is typically asymptomatic yet is most readily treated. Walter J. Burdette, Cancer: Etiology, Diagnosis, and Treatment 166 (1998); Memarzadeh & Berek, *supra*; Runnebaum & Stickeler, *supra*; Werness & Eltabbakh, *supra*. Commonly used screening tests include biannual rectovaginal pelvic examination,

15   radioimmunoassay to detect the CA-125 serum tumor marker, and transvaginal ultrasonography. Burdette, *supra* at 166.

Pelvic examination has failed to yield adequate numbers of early diagnoses, and the other methods are not sufficiently accurate. *Id.* One study reported that only 15% of patients who suffered from ovarian cancer were diagnosed with the disease at the time of

20   their pelvic examination. Look, *supra* at 174. Moreover, the CA-125 test is prone to giving false positives in pre-menopausal women and has been reported to be of low predictive value in post-menopausal women. *Id.* at 174-75. Although transvaginal ultrasonography is now the preferred procedure for screening for ovarian cancer, it is unable to distinguish reliably between benign and malignant tumors, and also cannot

25   locate primary peritoneal malignancies or ovarian cancer if the ovary size is normal. Schilder et al., *supra* at 194-95. While genetic testing for mutations of the BRCA1, BRCA2, hMSH2, and hMLH1 genes is now available, these tests may be too costly for some patients and may also yield false negative or indeterminate results. Schilder et al., *supra* at 191-94.

30   The staging of ovarian cancer, which is accomplished through surgical exploration, is crucial in determining the course of treatment and management of the disease. AJCC Cancer Staging Handbook 187 (Irvin D. Fleming et al. eds., 5th ed. 1998); Burdette, *supra* at 170; Memarzadeh & Berek, *supra*; Shridhar et al., *supra*. Staging is performed

by reference to the classification system developed by the International Federation of Gynecology and Obstetrics. David H. Moore, *Primary Surgical Management of Early Epithelial Ovarian Carcinoma*, in Ovarian Cancer 203 (Stephen C. Rubin & Gregory P. Sutton eds., 2d ed. 2001); Fleming et al. eds., *supra* at 188. Stage I ovarian cancer is

5 characterized by tumor growth that is limited to the ovaries and is comprised of three substages. *Id.* In substage IA, tumor growth is limited to one ovary, there is no tumor on .the external surface of the ovary, the ovarian capsule is intact, and no malignant cells are present in ascites or peritoneal washings. *Id.* Substage IB is identical to A1, except that tumor growth is limited to both ovaries. *Id.* Substage IC refers to the presence of tumor

10 growth limited to one or both ovaries, and also includes one or more of the following characteristics: capsule rupture, tumor growth on the surface of one or both ovaries, and malignant cells present in ascites or peritoneal washings. *Id.*

Stage II ovarian cancer refers to tumor growth involving one or both ovaries, along with pelvic extension. *Id.* Substage IIA involves extension and/or implants on the uterus

15 and/or fallopian tubes, with no malignant cells in the ascites or peritoneal washings, while substage IIB involves extension into other pelvic organs and tissues, again with no malignant cells in the ascites or peritoneal washings. *Id.* Substage IIC involves pelvic extension as in IIA or IIB, but with malignant cells in the ascites or peritoneal washings. *Id.*

20 Stage III ovarian cancer involves tumor growth in one or both ovaries, with peritoneal metastasis beyond the pelvis confirmed by microscope and/or metastasis in the regional lymph nodes. *Id.* Substage IIIA is characterized by microscopic peritoneal metastasis outside the pelvis, with substage IIIB involving macroscopic peritoneal metastasis outside the pelvis 2 cm or less in greatest dimension. *Id.* Substage IIIC is

25 identical to IIIB, except that the metastasis is greater than 2 cm in greatest dimension and may include regional lymph node metastasis. *Id.* Lastly, Stage IV refers to the presence distant metastasis, excluding peritoneal metastasis. *Id.*

While surgical staging is currently the benchmark for assessing the management and treatment of ovarian cancer, it suffers from considerable drawbacks, including the

30 invasiveness of the procedure, the potential for complications, as well as the potential for inaccuracy. Moore, *supra* at 206-208, 213. In view of these limitations, attention has turned to developing alternative staging methodologies through understanding differential gene expression in various stages of ovarian cancer and by obtaining various biomarkers

5

to help better assess the progression of the disease. Vartiainen, J. et al., *Int'l J. Cancer*, 95(5): 313-16 (2001); Shridhar et al. *supra*; Baekelandt, M. et al., *J. Clin. Oncol.* 18(22): 3775-81.

5   The treatment of ovarian cancer typically involves a multiprong attack, with surgical intervention serving as the foundation of treatment. Dennis S. Chi & William J. Hoskins, *Primary Surgical Management of Advanced Epithelial Ovarian Cancer*, in <u>Ovarian Cancer</u> 241 (Stephen C. Rubin & Gregory P. Sutton eds., 2d ed. 2001). For example, in the case of epithelial ovarian cancer, which accounts for ~90% of cases of ovarian cancer, treatment typically consists of: (1) cytoreductive surgery, including total

10   abdominal hysterectomy, bilateral salpingo-oophorectomy, omentectomy, and lymphadenectomy, followed by (2) adjuvant chemotherapy with paclitaxel and either cisplatin or carboplatin. Eltabbakh, G.H. & Awtrey, C.S., *Expert Op. Pharmacother.* 2(10): 109-24. Despite a clinical response rate of 80% to the adjuvant therapy, most patients experience tumor recurrence within three years of treatment. *Id.* Certain patients

15   may undergo a second cytoreductive surgery and/or second-line chemotherapy. Memarzadeh & Berek, *supra*.

From the foregoing, it is clear that procedures used for detecting, diagnosing, monitoring, staging, prognosticating, and preventing the recurrence of ovarian cancer are of critical importance to the outcome of the patient. Moreover, current procedures, while

20   helpful in each of these analyses, are limited by their specificity, sensitivity, invasiveness, and/or their cost. As such, highly specific and sensitive procedures that would operate by way of detecting novel markers in cells, tissues, or bodily fluids, with minimal invasiveness and at a reasonable cost, would be highly desirable.

Accordingly, there is a great need for more sensitive and accurate methods for

25   predicting whether a person is likely to develop ovarian cancer, for diagnosing ovarian cancer, for monitoring the progression of the disease, for staging the ovarian cancer, for determining whether the ovarian cancer has metastasized, and for imaging the ovarian cancer. There is also a need for better treatment of ovarian cancer.

30   SUMMARY OF THE INVENTION

The present invention solves many needs in the art by providing nucleic acid molecules, polypeptides and antibodies thereto, variants and derivatives of the nucleic acids and polypeptides, agonists and antagonists that may be used to identify, diagnose,

monitor, stage, image and treat ovarian cancer and non-cancerous disease states in ovarian; identify and monitor ovarian tissue; and identify and design agonists and antagonists of polypeptides of the invention. The invention also provides gene therapy, methods for producing transgenic animals and cells, and methods for producing

5   engineered ovarian tissue for treatment and research.

One aspect of the present invention relates to nucleic acid molecules that are specific to ovarian cells, ovarian tissue and/or the ovarian organ. These ovarian specific nucleic acids (OSNAs) may be a naturally occurring cDNA, genomic DNA, RNA, or a fragment of one of these nucleic acids, or may be a non-naturally occurring nucleic acid

10   molecule. If the OSNA is genomic DNA, then the OSNA is a ovarian specific gene (OSG). If the OSNA is RNA, then it is a ovarian specific transcript encoded by a OSG. Due to alternative splicing and transcriptional modification one OSG may encode for multiple ovarian specific RNAs. In a preferred embodiment, the nucleic acid molecule encodes a polypeptide that is specific to ovarian. More preferred is a nucleic acid

15   molecule that encodes a polypeptide comprising an amino acid sequence of SEQ ID NO: 100-192. In another preferred embodiment, the nucleic acid molecule comprises a nucleic acid sequence of SEQ ID NO: 1-99. For the OSNA sequences listed herein, DEX0367_001.nt.1 corresponds to SEQ ID NO: 1. For sequences with multiple splice variants, the parent sequence DEX0367_001.nt.1, will be followed by DEX0367_001.nt.2,

20   etc. for each splice variant. The sequences off the corresponding peptides are listed as DEX0367_001.aa.1, etc. For the mapping of all of the nucleotides and peptides, see the table in the Example 1 section below.

This aspect of the present invention also relates to nucleic acid molecules that selectively hybridize or exhibit substantial sequence similarity to nucleic acid molecules

25   encoding a ovarian Specific Protein (OSP), or that selectively hybridize or exhibit substantial sequence similarity to a OSNA. In one embodiment of the present invention the nucleic acid molecule comprises an allelic variant of a nucleic acid molecule encoding a OSP, or an allelic variant of a OSNA. In another embodiment, the nucleic acid molecule comprises a part of a nucleic acid sequence that encodes a OSP or a part of a nucleic acid

30   sequence of a OSNA.

In addition, this aspect of the present invention relates to a nucleic acid molecule further comprising one or more expression control sequences controlling the transcription

and/or translation of all or a part of a OSNA or the transcription and/or translation of a nucleic acid molecule that encodes all or a fragment of a OSP.

Another aspect of the present invention relates to vectors and/or host cells comprising a nucleic acid molecule of this invention. In a preferred embodiment, the

5   nucleic acid molecule of the vector and/or host cell encodes all or a fragment of a OSP. In another preferred embodiment, the nucleic acid molecule of the vector and/or host cell comprises all or a part of a OSNA. Vectors and host cells of the present invention are useful in the recombinant production of polypeptides, particularly OSPs of the present invention.

10   Another aspect of the present invention relates to polypeptides encoded by a nucleic acid molecule of this invention. The polypeptide may comprise either a fragment or a full-length protein. In a preferred embodiment, the polypeptide is a OSP. However, this aspect of the present invention also relates to mutant proteins (muteins) of OSPs, fusion proteins of which a portion is a OSP, and proteins and polypeptides encoded by

15   allelic variants of a OSNA as provided herein.

Another aspect of the present invention relates to antibodies and other binders that specifically binds to a polypeptide of the instant invention. Accordingly antibodies or binders of the present specifically bind to OSPs, muteins, fusion proteins, and/or homologous proteins or a polypeptides encoded by allelic variants of an OSNA as

20   provided herein.

Another aspect of the present invention relates to agonists and antagonists of the nucleic acid molecules and polypeptides of this invention. The agonists and antagonists of the instant invention may be used to treat ovarian cancer and non-cancerous disease states in ovarian and to produce engineered ovarian tissue.

25   Another aspect of the present invention relates to methods for using the nucleic acid molecules to detect or amplify nucleic acid molecules that have similar or identical nucleic acid sequences compared to the nucleic acid molecules described herein. Such methods are useful in identifying, diagnosing, monitoring, staging, imaging and treating ovarian cancer and non-cancerous disease states in ovarian. Such methods are also useful

30   in identifying and/or monitoring ovarian tissue. In addition, measurement of levels of one or more of the nucleic acid molecules of this invention may be useful for diagnostics as part of panel in combination with known other markers, particularly those described in the ovarian cancer background section above.

Another aspect of the present invention relates to use of the nucleic acid molecules of this invention in gene therapy, for producing transgenic animals and cells, and for producing engineered ovarian tissue for treatment and research.

Another aspect of the present invention relates to methods for detecting

5   polypeptides this invention, preferably using antibodies thereto. Such methods are useful to identify, diagnose, monitor, stage, image and treat ovarian cancer and non-cancerous disease states in ovarian. In addition, measurement of levels of one or more of the polypeptides of this invention may be useful to identify, diagnose, monitor, stage, image ovarian cancer in combination with known other markers, particularly those described in

10   the ovarian cancer background section above. The polypeptides of the present invention can also be used to identify and/or monitor ovarian tissue, and to produce engineered ovarian tissue.

Yet another aspect of the present invention relates to a computer readable means of storing the nucleic acid and amino acid sequences of the invention. The records of the

15   computer readable means can be accessed for reading and displaying of sequences for comparison, alignment and ordering of the sequences of the invention to other sequences. In addition, the computer records regarding the nucleic acid and/or amino acid sequences and/or measurements of their levels may be used alone or in combination with other markers to diagnose ovarian related diseases.

20   **DETAILED DESCRIPTION OF THE INVENTION**

Definitions and General Techniques

Unless otherwise defined herein, scientific and technical terms used in connection with the present invention shall have the meanings that are commonly understood by those of ordinary skill in the art. Further, unless otherwise required by context, singular terms

25   shall include pluralities and plural terms shall include the singular. Generally, nomenclatures used in connection with, and techniques of, cell and tissue culture, molecular biology, immunology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well known and commonly used in the art. The methods and techniques of the present invention are generally performed

30   according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification unless otherwise indicated. *See, e.g.*, Sambrook *et al.*, Molecular Cloning: A

Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press (1989) and Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 3d ed., Cold Spring Harbor Press (2001); Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing Associates (1992, and Supplements to 2000); Ausubel *et al.*, Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology – 4th Ed., Wiley & Sons (1999); Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1990); and Harlow and Lane, Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1999).

Enzymatic reactions and purification techniques are performed according to manufacturer's specifications, as commonly accomplished in the art or as described herein. The nomenclatures used in connection with, and the laboratory procedures and techniques of, analytical chemistry, synthetic organic chemistry, and medicinal and pharmaceutical chemistry described herein are those well known and commonly used in the art. Standard techniques are used for chemical syntheses, chemical analyses, pharmaceutical preparation, formulation, and delivery, and treatment of patients.

The following terms, unless otherwise indicated, shall be understood to have the following meanings:

A "nucleic acid molecule" of this invention refers to a polymeric form of nucleotides and includes both sense and antisense strands of RNA, cDNA, genomic DNA, and synthetic forms and mixed polymers of the above. A nucleotide refers to a ribonucleotide, deoxynucleotide or a modified form of either type of nucleotide. A "nucleic acid molecule" as used herein is synonymous with "nucleic acid" and "polynucleotide." The term "nucleic acid molecule" usually refers to a molecule of at least 10 bases in length, unless otherwise specified. The term includes single and double stranded forms of DNA. In addition, a polynucleotide may include either or both naturally occurring and modified nucleotides linked together by naturally occurring and/or non-naturally occurring nucleotide linkages.

Nucleotides are represented by single letter symbols in nucleic acid molecule sequences. The following table lists symbols identifying nucleotides or groups of nucleotides which may occupy the symbol position on a nucleic acid molecule. *See* Nomenclature Committee of the International Union of Biochemistry (NC-IUB), Nomenclature for incompletely specified bases in nucleic acid sequences, Recommendations 1984., *Eur J Biochem.* 150(1):1-5 (1985).

Symbol	Meaning	Group/Origin of Designation	Complementary Symbol
a	a	Adenine	t/u
g	g	Guanine	c
c	c	Cytosine	g
t	t	Thymine	a
u	u	Uracil	a
r	g or a	puRine	y
y	t/u or c	pYrimidine	r
m	a or c	aMino	k
k	g or t/u	Keto	m
s	g or c	Strong interactions 3H-bonds	w
w	a or t/u	Weak interactions 2H-bonds	s
b	g or c or t/u	not a	v
d	a or g or t/u	not c	h
h	a or c or t/u	not g	d
v	a or g or c	not t, not u	b
n	a or g or c or t/u, unknown, or other	aNy	n

The nucleic acid molecules may be modified chemically or biochemically or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those of skill in the art. Such modifications include, for example, labels, methylation,

5   substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoramidates, carbamates, etc.), charged linkages (*e.g.*, phosphorothioates, phosphorodithioates, etc.), pendent moieties (*e.g.*, polypeptides), intercalators (*e.g.*, acridine, psoralen, etc.), chelators, alkylators, and modified linkages

10   (*e.g.*, alpha anomeric nucleic acids, etc.) The term "nucleic acid molecule" also includes any topological conformation, including single-stranded, double-stranded, partially duplexed, triplexed, hairpinned, circular and padlocked conformations. Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Such molecules are

15   known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of the molecule.

A "gene" is defined as a nucleic acid molecule that comprises a nucleic acid sequence that encodes a polypeptide and the expression control sequences that surround the nucleic acid sequence that encodes the polypeptide. For instance, a gene may

20   comprise a promoter, one or more enhancers, a nucleic acid sequence that encodes a polypeptide, downstream regulatory sequences and, possibly, other nucleic acid sequences

involved in regulation of the expression of an RNA. As is well known in the art, eukaryotic genes usually contain both exons and introns. The term "exon" refers to a nucleic acid sequence found in genomic DNA that is bioinformatically predicted and/or experimentally confirmed to contribute contiguous sequence to a mature mRNA

5   transcript. The term "intron" refers to a nucleic acid sequence found in genomic DNA that is predicted and/or confirmed to not contribute to a mature mRNA transcript, but rather to be "spliced out" during processing of the transcript.

A nucleic acid molecule or polypeptide is "derived" from a particular species if the nucleic acid molecule or polypeptide has been isolated from the particular species, or if the

10   nucleic acid molecule or polypeptide is homologous to a nucleic acid molecule or polypeptide isolated from a particular species.

An "isolated" or "substantially pure" nucleic acid or polynucleotide (*e.g.*, an RNA, DNA or a mixed polymer) is one which is substantially separated from other cellular components that naturally accompany the native polynucleotide in its natural host cell,

15   *e.g.*, ribosomes, polymerases, or genomic sequences with which it is naturally associated. The term embraces a nucleic acid or polynucleotide that (1) has been removed from its naturally occurring environment, (2) is not associated with all or a portion of a polynucleotide in which the "isolated polynucleotide" is found in nature, (3) is operatively linked to a polynucleotide which it is not linked to in nature, (4) does not occur in nature

20   as part of a larger sequence or (5) includes nucleotides or internucleoside bonds that are not found in nature. The term "isolated" or "substantially pure" also can be used in reference to recombinant or cloned DNA isolates, chemically synthesized polynucleotide analogs, or polynucleotide analogs that are biologically synthesized by heterologous systems. The term "isolated nucleic acid molecule" includes nucleic acid molecules that

25   are integrated into a host cell chromosome at a heterologous site, recombinant fusions of a native fragment to a heterologous sequence, recombinant vectors present as episomes or as integrated into a host cell chromosome.

A "part" of a nucleic acid molecule refers to a nucleic acid molecule that comprises a partial contiguous sequence of at least 10 bases of the reference nucleic acid

30   molecule. Preferably, a part comprises at least 15 to 20 bases of a reference nucleic acid molecule. In theory, a nucleic acid sequence of 17 nucleotides is of sufficient length to occur at random less frequently than once in the three gigabase human genome, and thus to provide a nucleic acid probe that can uniquely identify the reference sequence in a

nucleic acid mixture of genomic complexity. A preferred part is one that comprises a nucleic acid sequence that can encode at least 6 contiguous amino acid sequences (fragments of at least 18 nucleotides) because they are useful in directing the expression or synthesis of peptides that are useful in mapping the epitopes of the polypeptide encoded

5   by the reference nucleic acid. *See, e.g.*, Geysen *et al.*, *Proc. Natl. Acad. Sci. USA* 81:3998-4002 (1984); and U.S. Patent Nos. 4,708,871 and 5,595,915, the disclosures of which are incorporated herein by reference in their entireties. A part may also comprise at least 25, 30, 35 or 40 nucleotides of a reference nucleic acid molecule, or at least 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400 or 500 nucleotides of a reference nucleic

10   acid molecule. A part of a nucleic acid molecule may comprise no other nucleic acid sequences. Alternatively, a part of a nucleic acid may comprise other nucleic acid sequences from other nucleic acid molecules.

The term "oligonucleotide" refers to a nucleic acid molecule generally comprising a length of 200 bases or fewer. The term often refers to single-stranded

15   deoxyribonucleotides, but it can refer as well to single-or double-stranded ribonucleotides, RNA:DNA hybrids and double-stranded DNAs, among others. Preferably, oligonucleotides are 10 to 60 bases in length and most preferably 12, 13, 14, 15, 16, 17, 18, 19 or 20 bases in length. Other preferred oligonucleotides are 25, 30, 35, 40, 45, 50, 55 or 60 bases in length. Oligonucleotides may be single-stranded, *e.g.* for use as probes

20   or primers, or may be double-stranded, *e.g.* for use in the construction of a mutant gene. Oligonucleotides of the invention can be either sense or antisense oligonucleotides. An oligonucleotide can be derivatized or modified as discussed above for nucleic acid molecules.

Oligonucleotides, such as single-stranded DNA probe oligonucleotides, often are

25   synthesized by chemical methods, such as those implemented on automated oligonucleotide synthesizers. However, oligonucleotides can be made by a variety of other methods, including in vitro recombinant DNA-mediated techniques and by expression of DNAs in cells and organisms. Initially, chemically synthesized DNAs typically are obtained without a 5' phosphate. The 5' ends of such oligonucleotides are

30   not substrates for phosphodiester bond formation by ligation reactions that employ DNA ligases typically used to form recombinant DNA molecules. Where ligation of such oligonucleotides is desired, a phosphate can be added by standard techniques, such as those that employ a kinase and ATP. The 3' end of a chemically synthesized

oligonucleotide generally has a free hydroxyl group and, in the presence of a ligase, such as T4 DNA ligase, readily will form a phosphodiester bond with a 5' phosphate of another polynucleotide, such as another oligonucleotide. As is well known, this reaction can be prevented selectively, where desired, by removing the 5' phosphates of the other

5   polynucleotide(s) prior to ligation.

The term "naturally occurring nucleotide" referred to herein includes naturally occurring deoxyribonucleotides and ribonucleotides. The term "modified nucleotides" referred to herein includes nucleotides with modified or substituted sugar groups and the like. The term "nucleotide linkages" referred to herein includes nucleotides linkages such

10   as phosphorothioate, phosphorodithioate, phosphoroselenoate, phosphorodiselenoate, phosphoroanilothioate, phoshoraniladate, phosphoroamidate, and the like. *See e.g.,* LaPlanche *et al. Nucl. Acids Res.* 14:9081-9093 (1986); Stein *et al. Nucl. Acids Res.* 16:3209-3221 (1988); Zon *et al. Anti-Cancer Drug Design* 6:539-568 (1991); Zon *et al.,* in Eckstein (ed.) Oligonucleotides and Analogues: A Practical Approach, pp. 87-108,

15   Oxford University Press (1991); Uhlmann and Peyman *Chemical Reviews* 90:543 (1990), and U.S. Patent No. 5,151,510, the disclosure of which is hereby incorporated by reference in its entirety.

Unless specified otherwise, the left hand end of a polynucleotide sequence in sense orientation is the 5' end and the right hand end of the sequence is the 3' end. In addition,

20   the left hand direction of a polynucleotide sequence in sense orientation is referred to as the 5' direction, while the right hand direction of the polynucleotide sequence is referred to as the 3' direction. Further, unless otherwise indicated, each nucleotide sequence is set forth herein as a sequence of deoxyribonucleotides. It is intended, however, that the given sequence be interpreted as would be appropriate to the polynucleotide composition: for

25   example, if the isolated nucleic acid is composed of RNA, the given sequence intends ribonucleotides, with uridine substituted for thymidine.

The term "allelic variant" refers to one of two or more alternative naturally occurring forms of a gene, wherein each gene possesses a unique nucleotide sequence. In a preferred embodiment, different alleles of a given gene have similar or identical

30   biological properties.

The term "percent sequence identity" in the context of nucleic acid sequences refers to the residues in two sequences which are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over a stretch of at

14

least about nine nucleotides, usually at least about 20 nucleotides, more usually at least
about 24 nucleotides, typically at least about 28 nucleotides, more typically at least about
32 nucleotides, and preferably at least about 36 or more nucleotides. There are a number
of different algorithms known in the art which can be used to measure nucleotide sequence
5   identity. For instance, polynucleotide sequences can be compared using FASTA, Gap or
Bestfit, which are programs in Wisconsin Package Version 10.0, Genetics Computer
Group (GCG), Madison, Wisconsin. FASTA, which includes, *e.g.*, the programs FASTA2
and FASTA3, provides alignments and percent sequence identity of the regions of the best
overlap between the query and search sequences (Pearson, *Methods Enzymol.* 183: 63-98
10   (1990); Pearson, *Methods Mol. Biol.* 132: 185-219 (2000); Pearson, *Methods Enzymol.*
266: 227-258 (1996); Pearson, *J. Mol. Biol.* 276: 71-84 (1998)). Unless otherwise
specified, default parameters for a particular program or algorithm are used. For instance, ·
· percent sequence identity between nucleic acid sequences can be determined using
FASTA with its default parameters (a word size of 6 and the NOPAM factor for the
15   scoring matrix) or using Gap with its default parameters as provided in GCG Version 6.1. ··

    A reference to a nucleic acid sequence encompasses its complement unless
otherwise specified. Thus, a reference to a nucleic acid molecule having a particular
sequence should be understood to encompass its complementary strand, with its
complementary sequence. The complementary strand is also useful, *e.g.*, for antisense
20   therapy, double stranded RNA (dsRNA) inhibition (RNAi), combination of triplex and
antisense, hybridization probes and PCR primers.

    In the molecular biology art, researchers use the terms "percent sequence identity",
"percent sequence similarity" and "percent sequence homology" interchangeably. In this
application, these terms shall have the same meaning with respect to nucleic acid
25   sequences only.

    The term "substantial similarity" or "substantial sequence similarity," when
referring to a nucleic acid or fragment thereof, indicates that, when optimally aligned with
appropriate nucleotide insertions or deletions with another nucleic acid (or its
complementary strand), there is nucleotide sequence identity in at least about 50%, more
30   preferably 60% of the nucleotide bases, usually at least about 70%, more usually at least
about 80%, preferably at least about 90%, and more preferably at least about 95-98% of
the nucleotide bases, as measured by any well known algorithm of sequence identity, such
as FASTA, BLAST or Gap, as discussed above.

Alternatively, substantial similarity exists between a first and second nucleic acid sequence when the first nucleic acid sequence or fragment thereof hybridizes to an antisense strand of the second nucleic acid, under selective hybridization conditions. Typically, selective hybridization will occur between the first nucleic acid sequence and

5   an antisense strand of the second nucleic acid sequence when there is at least about 55% sequence identity between the first and second nucleic acid sequences— preferably at least about 65%, more preferably at least about 75%, and most preferably at least about 90% — over a stretch of at least about 14 nucleotides, more preferably at least 17 nucleotides, even more preferably at least 20, 25, 30, 35, 40, 50, 60, 70, 80, 90 or 100 nucleotides.

10   Nucleic acid hybridization will be affected by such conditions as salt concentration, temperature, solvents, the base composition of the hybridizing species, length of the complementary regions, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be readily appreciated by those skilled in the art. "Stringent hybridization conditions" and "stringent wash conditions" in the context

15   of nucleic acid hybridization experiments depend upon a number of different physical parameters. The most important parameters include temperature of hybridization, base composition of the nucleic acids, salt concentration and length of the nucleic acid. One having ordinary skill in the art knows how to vary these parameters to achieve a particular stringency of hybridization. In general, "stringent hybridization" is performed at about

20   25°C below the thermal melting point (T_m) for the specific DNA hybrid under a particular set of conditions. "Stringent washing" is performed at temperatures about 5°C lower than the T_m for the specific DNA hybrid under a particular set of conditions. The T_m is the temperature at which 50% of the target sequence hybridizes to a perfectly matched probe. *See* Sambrook (1989), *supra*, p. 9.51.

25   The T_m for a particular DNA-DNA hybrid can be estimated by the formula:

$$T_m = 81.5°C + 16.6 (\log_{10}[Na^+]) + 0.41 (\text{fraction } G + C) -$$
0.63 (% formamide) - (600/l) where l is the length of the hybrid in base pairs.

The T_m for a particular RNA-RNA hybrid can be estimated by the formula:

$$T_m = 79.8°C + 18.5 (\log_{10}[Na^+]) + 0.58 (\text{fraction } G + C) +$$

30    $$11.8 (\text{fraction } G + C)^2 - 0.35 (\% \text{ formamide}) - (820/l).$$

The T_m for a particular RNA-DNA hybrid can be estimated by the formula:

$$T_m = 79.8°C + 18.5(\log_{10}[Na^+]) + 0.58 (\text{fraction } G + C) +$$
$$11.8 (\text{fraction } G + C)^2 - 0.50 (\% \text{ formamide}) - (820/l).$$

In general, the T_m decreases by 1-1.5°C for each 1% of mismatch between two nucleic acid sequences. Thus, one having ordinary skill in the art can alter hybridization and/or washing conditions to obtain sequences that have higher or lower degrees of sequence identity to the target nucleic acid. For instance, to obtain hybridizing nucleic

5　acids that contain up to 10% mismatch from the target nucleic acid sequence, 10-15°C would be subtracted from the calculated T_m of a perfectly matched hybrid, and then the hybridization and washing temperatures adjusted accordingly. Probe sequences may also hybridize specifically to duplex DNA under certain conditions to form triplex or other higher order DNA complexes. The preparation of such probes and suitable hybridization

10　conditions are well known in the art.

An example of stringent hybridization conditions for hybridization of complementary nucleic acid sequences having more than 100 complementary residues on a filter in a Southern or Northern blot or for screening a library is 50% formamide/6X SSC at 42°C for at least ten hours and preferably overnight (approximately 16 hours). Another

15　example of stringent hybridization conditions is 6X SSC at 68°C without formamide for at least ten hours and preferably overnight. An example of moderate stringency hybridization conditions is 6X SSC at 55°C without formamide for at least ten hours and preferably overnight. An example of low stringency hybridization conditions for hybridization of complementary nucleic acid sequences having more than 100

20　complementary residues on a filter in a Southern or northern blot or for screening a library is 6X SSC at 42°C for at least ten hours. Hybridization conditions to identify nucleic acid sequences that are similar but not identical can be identified by experimentally changing the hybridization temperature from 68°C to 42°C while keeping the salt concentration constant (6X SSC), or keeping the hybridization temperature and salt concentration

25　constant (*e.g.* 42°C and 6X SSC) and varying the formamide concentration from 50% to 0%. Hybridization buffers may also include blocking agents to lower background. These agents are well known in the art. *See* Sambrook *et al.* (1989), *supra*, pages 8.46 and 9.46-9.58. *See also* Ausubel (1992), *supra*, Ausubel (1999), *supra*, and Sambrook (2001), *supra.*

30　　　　Wash conditions also can be altered to change stringency conditions. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (*see* Sambrook (1989), *supra*, for SSC buffer). Often the high stringency wash is preceded by a low stringency wash to remove excess probe. An exemplary medium stringency wash for

duplex DNA of more than 100 base pairs is 1x SSC at 45°C for 15 minutes. An exemplary low stringency wash for such a duplex is 4x SSC at 40°C for 15 minutes. In general, signal-to-noise ratio of 2x or higher than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization.

5   As defined herein, nucleic acids that do not hybridize to each other under stringent conditions are still substantially similar to one another if they encode polypeptides that are substantially identical to each other. This occurs, for example, when a nucleic acid is created synthetically or recombinantly using a high codon degeneracy as permitted by the redundancy of the genetic code.

10   .Hybridization conditions for nucleic acid molecules that are shorter than 100 nucleotides in length (e.g., for oligonucleotide probes) may be calculated by the formula:

$$T_m = 81.5°C + 16.6(\log_{10}[Na^+]) + 0.41(\text{fraction G+C}) -(600/N),$$ wherein N is change length and the $[Na^+]$ is 1 M or less. See Sambrook (1989), supra, p. 11.46. For hybridization of probes shorter than 100 nucleotides, hybridization is usually performed

15 under stringent conditions (5-10°C below the T_m) using high concentrations (0.1-1.0 pmol/ml) of probe. Id. at p. 11.45. Determination of hybridization using mismatched probes, pools of degenerate probes or "guessmers," as well as hybridization solutions and methods for empirically determining hybridization conditions are well known in the art. See, e.g., Ausubel (1999), supra; Sambrook (1989), supra, pp. 11.45-11.57.

20   The term "digestion" or "digestion of DNA" refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes referred to herein are commercially available and their reaction conditions, cofactors and other requirements for use are known and routine to the skilled artisan. For analytical purposes, typically, 1 μg of plasmid or DNA fragment is

25 digested with about 2 units of enzyme in about 20 μl of reaction buffer. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μg of DNA are digested with 20 to 250 units of enzyme in proportionately larger volumes. Appropriate buffers and substrate amounts for particular restriction enzymes are described in standard laboratory manuals, such as those referenced below, and are specified by commercial

30 suppliers. Incubation times of about 1 hour at 37°C are ordinarily used, but conditions may vary in accordance with standard procedures, the supplier's instructions and the particulars of the reaction. After digestion, reactions may be analyzed, and fragments may

be purified by electrophoresis through an agarose or polyacrylamide gel, using well
known methods that are routine for those skilled in the art.

The term "ligation" refers to the process of forming phosphodiester bonds between
two or more polynucleotides, which most often are double-stranded DNAs. Techniques

5　for ligation are well known to the art and protocols for ligation are described in standard
laboratory manuals and references, such as, *e.g.*, Sambrook (1989), *supra.*

Genome-derived "single exon probes," are probes that comprise at least part of an
exon ("reference exon") and can hybridize detectably under high stringency conditions to
transcript-derived nucleic acids that include the reference exon but do not hybridize

10　detectably under high stringency conditions to nucleic acids that lack the reference exon.
Single exon probes typically further comprise, contiguous to a first end of the exon
portion, a first intronic and/or intergenic sequence that is identically contiguous to the
exon in the genome, and may contain a second intronic and/or intergenic sequence that is
identically contiguous to the exon in the genome. The minimum length of genome-

15　derived single exon probes is defined by the requirement that the exonic portion be of
sufficient length to hybridize under high stringency conditions to transcript-derived
nucleic acids, as discussed above. The maximum length of genome-derived single exon
probes is defined by the requirement that the probes contain portions of no more than one
exon. The single exon probes may contain priming sequences not found in contiguity with

20　the rest of the probe sequence in the genome, which priming sequences are useful for PCR
and other amplification-based technologies. In another aspect, the invention is directed to
single exon probes based on the OSNAs disclosed herein.

In one embodiment, the term "microarray" refers to a "nucleic acid microarray"
having a substrate-bound plurality of nucleic acids, hybridization to each of the plurality

25　of bound nucleic acids being separately detectable. The substrate can be solid or porous,
planar or non-planar, unitary or distributed. Nucleic acid microarrays include all the
devices so called in Schena (ed.), DNA Microarrays: A Practical Approach (Practical
Approach Series), Oxford University Press (1999); *Nature Genet.* 21(1)(suppl.):1 - 60
(1999); Schena (ed.), Microarray Biochip: Tools and Technology, Eaton Publishing

30　Company/BioTechniques Books Division (2000). Additionally, these nucleic acid
microarrays include substrate-bound plurality of nucleic acids in which the plurality of
nucleic acids are disposed on a plurality of beads, rather than on a unitary planar substrate,
as is described, *inter alia*, in Brenner *et al.*, *Proc. Natl. Acad. Sci. USA* 97(4):1665-1670

19

(2000). Examples of nucleic acid microarrays may be found in U.S. Patent Nos. 6,391,623, 6,383,754, 6,383,749, 6,380,377, 6,379,897, 6,376,191, 6,372,431, 6,351,712 6,344,316, 6,316,193, 6,312,906, 6,309,828, 6,309,824, 6,306,643, 6,300,063, 6,287,850, 6,284,497, 6,284,465, 6,280,954, 6,262,216, 6,251,601, 6,245,518, 6,263,287, 6,251,601,

5   6,238,866, 6,228,575, 6,214,587, 6,203,989, 6,171,797, 6,103,474, 6,083,726, 6,054,274, 6,040,138, 6,083,726, 6,004,755, 6,001,309, 5,958,342, 5,952,180, 5,936,731, 5,843,655, 5,814,454, 5,837,196, 5,436,327, 5,412,087, 5,405,783, the disclosures of which are incorporated herein by reference in their entireties.

In an alternative embodiment, a "microarray" may also refer to a "peptide

10   microarray" or "protein microarray" having a substrate-bound collection of plurality of polypeptides, the binding to each of the plurality of bound polypeptides being separately detectable. Alternatively, the peptide microarray ,may have a plurality of binders, including but not limited to monoclonal antibodies, polyclonal antibodies, phage display binders, yeast 2 hybrid binders, aptamers, which can specifically detect the binding of the

15   polypeptides of this invention. The array may be based on autoantibody detection to the polypeptides of this invention, see Robinson *et al., Nature Medicine* 8(3):295-301 (2002). Examples of peptide arrays may be found in WO 02/31463, WO 02/25288, WO 01/94946, WO 01/88162, WO 01/68671, WO 01/57259, WO 00/61806, WO 00/54046, WO 00/47774, WO 99/40434, WO 99/39210, WO 97/42507 and U.S. Patent Nos. 6,268,210,

20   5,766,960, 5,143,854, the disclosures of which are incorporated herein by reference in their entireties.

In addition, determination of the levels of the OSNA or OSP may be made in a multiplex manner using techniques described in WO 02/29109, WO 02/24959, WO 01/83502, WO01/73113, WO 01/59432, WO 01/57269, WO 99/67641, the disclosures of

25   which are incorporated herein by reference in their entireties.

The term "mutant", "mutated", or "mutation" when applied to nucleic acid sequences means that nucleotides in a nucleic acid sequence may be inserted, deleted or changed compared to a reference nucleic acid sequence. A single alteration may be made at a locus (a point mutation) or multiple nucleotides may be inserted, deleted or changed at

30   a single locus. In addition, one or more alterations may be made at any number of loci within a nucleic acid sequence. In a preferred embodiment of the present invention, the nucleic acid sequence is the wild type nucleic acid sequence encoding a OSP or is a

20

OSNA. The nucleic acid sequence may be mutated by any method known in the art including those mutagenesis techniques described *infra*.

The term "error-prone PCR" refers to a process for performing PCR under conditions where the copying fidelity of the DNA polymerase is low, such that a high rate

5   of point mutations is obtained along the entire length of the PCR product. *See, e.g.*, Leung *et al., Technique* 1: 11-15 (1989) and Caldwell *et al., PCR Methods Applic.* 2: 28-33 (1992).

The term "oligonucleotide-directed mutagenesis" refers to a process which enables the generation of site-specific mutations in any cloned DNA segment of interest. *See, e.g.*,

10   Reidhaar-Olson *et al., Science* 241: 53-57 (1988).

The term "assembly PCR" refers to a process which involves the assembly of a PCR product from a mixture of small DNA fragments. A large number of different PCR reactions occur in parallel in the same vial, with the products of one reaction priming the products of another reaction.

15   The term "sexual PCR mutagenesis" or "DNA shuffling" refers to a method of error-prone PCR coupled with forced homologous recombination between DNA molecules of different but highly related DNA sequence *in vitro*, caused by random fragmentation of the DNA molecule based on sequence similarity, followed by fixation of · the crossover by primer extension in an error-prone PCR reaction. *See, e.g.*, Stemmer,

20   *Proc. Natl. Acad. Sci. U.S.A.* 91: 10747-10751 (1994). DNA shuffling can be carried out · between several related genes ("Family shuffling").

The term "*in vivo* mutagenesis" refers to a process of generating random mutations · in any cloned DNA of interest which involves the propagation of the DNA in a strain of bacteria such as *E. coli* that carries mutations in one or more of the DNA repair pathways. ·

25   These "mutator" strains have a higher random mutation rate than that of a wild-type parent. Propagating the DNA in a mutator strain will eventually generate random mutations within the DNA.

The term "cassette mutagenesis" refers to any process for replacing a small region of a double-stranded DNA molecule with a synthetic oligonucleotide "cassette" that

30   differs from the native sequence. The oligonucleotide often contains completely and/or partially randomized native sequence.

The term "recursive ensemble mutagenesis" refers to an algorithm for protein engineering (protein mutagenesis) developed to produce diverse populations of

phenotypically related mutants whose members differ in amino acid sequence. This method uses a feedback mechanism to control successive rounds of combinatorial cassette mutagenesis. *See, e.g.,* Arkin *et al., Proc. Natl. Acad. Sci. U.S.A.* 89: 7811-7815 (1992).

The term "exponential ensemble mutagenesis" refers to a process for generating

5   combinatorial libraries with a high percentage of unique and functional mutants, wherein small groups of residues are randomized in parallel to identify, at each altered position, amino acids which lead to functional proteins. *See, e.g.,* Delegrave *et al., Biotechnology Research* 11: 1548-1552 (1993); Arnold, *Current Opinion in Biotechnology* 4: 450-455 (1993).

10   "Operatively linked" expression control sequences refers to a linkage in which the expression control sequence is either contiguous with the gene of interest to control the gene of interest, or acts in *trans* or at a distance to control the gene of interest.

·   ·The term "expression control sequence" as used herein refers to polynucleotide sequences which are necessary to affect the expression of coding sequences to which they

15   are operatively linked. Expression control sequences are sequences which control the transcription, post-transcriptional events and translation of nucleic acid sequences. Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation signals; sequences that stabilize cytoplasmic mRNA; sequences that

20   enhance translation efficiency (*e.g.,* ribosome binding sites); sequences that enhance protein stability; and when desired, sequences that enhance protein secretion. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include promoter, ribosomal binding site, and transcription termination sequence. The term "control sequences" is intended to include, at a minimum,

25   all components whose presence is essential for expression, and can also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences.

The term "vector," as used herein, is intended to refer to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of

30   vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments may be ligated. Other vectors include cosmids, bacterial artificial chromosomes (BAC) and yeast artificial chromosomes (YAC). Another type of vector is a viral vector, wherein additional DNA segments may be ligated into the viral

genome. Viral vectors that infect bacterial cells are referred to as bacteriophages. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication). Other vectors can be integrated into the genome of a host cell upon introduction into the host cell, and thereby

5   are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "recombinant expression vectors" (or simply, "expression vectors"). In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" may be used

10   interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include other forms of expression vectors that serve equivalent functions.

The term "recombinant host cell" (or simply "host cell"), as used herein, is intended to refer to a cell into which a recombinant expression vector has been introduced.

15   It should be understood that such terms are intended to refer not only to the particular subject cell but to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term "host cell" as used herein.

20   As used herein, the phrase "open reading frame" and the equivalent acronym "ORF" refers to that portion of a transcript-derived nucleic acid that can be translated in its entirety into a sequence of contiguous amino acids. As so defined, an ORF has length, measured in nucleotides, exactly divisible by 3. As so defined, an ORF need not encode the entirety of a natural protein.

25   As used herein, the phrase "ORF-encoded peptide" refers to the predicted or actual translation of an ORF.

As used herein, the phrase "degenerate variant" of a reference nucleic acid sequence is meant to be inclusive of all nucleic acid sequences that can be directly translated, using the standard genetic code, to provide an amino acid sequence identical to

30   that translated from the reference nucleic acid sequence.

The term "polypeptide" encompasses both naturally occurring and non-naturally occurring proteins and polypeptides, as well as polypeptide fragments and polypeptide mutants, derivatives and analogs thereof. A polypeptide may be monomeric or polymeric.

Further, a polypeptide may comprise a number of different modules within a single polypeptide each of which has one or more distinct activities. A preferred polypeptide in accordance with the invention comprises a OSP encoded by a nucleic acid molecule of the instant invention, or a fragment, mutant, analog and derivative thereof.

5   The term "isolated protein" or "isolated polypeptide" is a protein or polypeptide that by virtue of its origin or source of derivation (1) is not associated with naturally associated components that accompany it in its native state, (2) is free of other proteins from the same species (3) is expressed by a cell from a different species, or (4) does not occur in nature. Thus, a polypeptide that is chemically synthesized or synthesized in a

10   cellular system different from the cell from which it naturally originates will be "isolated" from its naturally associated components. A polypeptide or protein may also be rendered substantially free of naturally associated components by isolation, using protein purification techniques well known in the art.

        A protein or polypeptide is "substantially pure," "substantially homogeneous" or

15   "substantially purified" when at least about 60% to 75% of a sample exhibits a single species of polypeptide. The polypeptide or protein may be monomeric or multimeric. A substantially pure polypeptide or protein will typically comprise about 50%, 60%, 70%, 80% or 90% W/W of a protein sample, more usually about 95%, and preferably will be over 99% pure. Protein purity or homogeneity may be determined by a number of means

20   well known in the art, such as polyacrylamide gel electrophoresis of a protein sample, followed by visualizing a single polypeptide band upon staining the gel with a stain well known in the art. For certain purposes, higher resolution may be provided by using HPLC or other means well known in the art for purification.

        The term "fragment" when used herein with respect to polypeptides of the present

25   invention refers to a polypeptide that has an amino-terminal and/or carboxy-terminal deletion compared to a full-length OSP. In a preferred embodiment, the fragment is a contiguous sequence in which the amino acid sequence of the fragment is identical to the corresponding positions in the naturally occurring polypeptide. Fragments typically are at least 5, 6, 7, 8, 9 or 10 amino acids long, preferably at least 12, 14, 16 or 18 amino acids

30   long, more preferably at least 20 amino acids long, more preferably at least 25, 30, 35, 40 or 45, amino acids, even more preferably at least 50 or 60 amino acids long, and even more preferably at least 70 amino acids long.

A "derivative" when used herein with respect to polypeptides of the present invention refers to a polypeptide which is substantially similar in primary structural sequence to a OSP but which include, *e.g.*, *in vivo* or *in vitro* chemical and biochemical modifications that are not found in the OSP. Such modifications include, for example,

5   acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate,

10   formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. Other modification include, *e.g.*, labeling with radionuclides, and various enzymatic

15   modifications, as will be readily appreciated by those skilled in the art. A variety of methods for labeling polypeptides and of substituents or labels useful for such purposes are well known in the art, and include radioactive isotopes such as ^{125}I, ^{32}P, ^{35}S, ^{14}C and ^3H, ligands which bind to labeled antiligands (*e.g.*, antibodies), fluorophores, chemiluminescent agents, enzymes, and antiligands which can serve as specific binding

20   pair members for a labeled ligand. The choice of label depends on the sensitivity required, ease of conjugation with the primer, stability requirements, and available instrumentation. Methods for labeling polypeptides are well known in the art. *See* Ausubel (1992), *supra;* Ausubel (1999), *supra.*

The term "fusion protein" refers to polypeptides of the present invention coupled

25   to a heterologous amino acid sequences. Fusion proteins are useful because they can be constructed to contain two or more desired functional elements from two or more different proteins. A fusion protein comprises at least 10 contiguous amino acids from a polypeptide of interest, more preferably at least 20 or 30 amino acids, even more preferably at least 40, 50 or 60 amino acids, yet more preferably at least 75, 100 or 125

30   amino acids. Fusion proteins can be produced recombinantly by constructing a nucleic acid sequence that encodes the polypeptide or a fragment thereof in frame with a nucleic acid sequence encoding a different protein or peptide and then expressing the fusion

protein. Alternatively, a fusion protein can be produced chemically by crosslinking the polypeptide or a fragment thereof to another protein.

The term "analog" refers to both polypeptide analogs and non-peptide analogs.

The term "polypeptide analog" as used herein refers to a polypeptide that is comprised of a

5   segment of at least 25 amino acids that has substantial identity to a portion of an amino acid sequence but which contains non-natural amino acids or non-natural inter-residue bonds. In a preferred embodiment, the analog has the same or similar biological activity as the native polypeptide. Typically, polypeptide analogs comprise a conservative amino acid substitution (or insertion or deletion) with respect to the naturally occurring sequence.

10   Analogs typically are at least 20 amino acids long, preferably at least 50 amino acids long or longer, and can often be as long as a full-length naturally occurring polypeptide.

The term "non-peptide analog" refers to a compound with properties that are analogous to those of a reference polypeptide. A non-peptide compound may also be termed a "peptide mimetic" or a "peptidomimetic." Such compounds are often developed

15   with the aid of computerized molecular modeling. Peptide mimetics that are structurally similar to useful peptides may be used to produce an equivalent effect. Generally, peptidomimetics are structurally similar to a paradigm polypeptide (*i.e.*, a polypeptide that has a desired biochemical property or pharmacological activity), but have one or more peptide linkages optionally replaced by a linkage selected from the group consisting of:

20   --CH$_2$NH--, --CH$_2$S--, --CH$_2$-CH$_2$--, --CH=CH--(cis and trans), --COCH$_2$--, --CH(OH)CH$_2$--, and --CH$_2$SO--, by methods well known in the art. Systematic substitution of one or more amino acids of a consensus sequence with a D-amino acid of the same type (*e.g.*, D-lysine in place of L-lysine) may also be used to generate more stable peptides. In addition, constrained peptides comprising a consensus sequence or a

25   substantially identical consensus sequence variation may be generated by methods known in the art (Rizo *et al., Ann. Rev. Biochem.* 61:387-418 (1992)). For example, one may add internal cysteine residues capable of forming intramolecular disulfide bridges which cyclize the peptide.

The term "mutant" or "mutein" when referring to a polypeptide of the present

30   invention relates to an amino acid sequence containing substitutions, insertions or deletions of one or more amino acids compared to the amino acid sequence of a OSP. A mutein may have one or more amino acid point substitutions, in which a single amino acid

at a position has been changed to another amino acid, one or more insertions and/or deletions, in which one or more amino acids are inserted or deleted, respectively, in the sequence of the naturally occurring protein, and/or truncations of the amino acid sequence at either or both the amino or carboxy termini. Further, a mutein may have the same or

5   different biological activity as the naturally occurring protein. For instance, a mutein may have an increased or decreased biological activity. A mutein has at least 50% sequence similarity to the wild type protein, preferred is 60% sequence similarity, more preferred is 70% sequence similarity. Even more preferred are muteins having 80%, 85% or 90% sequence similarity to a OSP. In an even more preferred embodiment, a mutein exhibits

10   95% sequence identity, even more preferably 97%, even more preferably 98% and even more preferably 99%. Sequence similarity may be measured by any common sequence analysis algorithm, such as GAP or BESTFIT or other variation Smith-Waterman alignment. *See*, T. F. Smith and M. S. Waterman, J. Mol. Biol. 147:195-197 (1981) and W.R. Pearson, Genomics 11:635-650 (1991).

15   Preferred amino acid substitutions are those which: (1) reduce susceptibility to proteolysis, (2) reduce susceptibility to oxidation, (3) alter binding affinity for forming protein complexes, (4) alter binding affinity or enzymatic activity, and (5) confer or modify other physicochemical or functional properties of such analogs. For example, · single or multiple amino acid substitutions (preferably conservative amino acid

20   substitutions) may be made in the naturally occurring sequence (preferably in the portion of the polypeptide outside the domain(s) forming intermolecular contacts. In a preferred embodiment, the amino acid substitutions are moderately conservative substitutions or conservative substitutions. In a more preferred embodiment, the amino acid substitutions are conservative substitutions. A conservative amino acid substitution should not

25   substantially change the structural characteristics of the parent sequence (*e.g.*, a replacement amino acid should not tend to disrupt a helix that occurs in the parent sequence, or disrupt other types of secondary structure that characterizes the parent sequence). Examples of art-recognized polypeptide secondary and tertiary structures are described in Creighton (ed.), Proteins, Structures and Molecular Principles, W. H.

30   Freeman and Company (1984); Branden *et al.* (ed.), Introduction to Protein Structure, Garland Publishing (1991); Thornton *et al.*, *Nature* 354:105-106 (1991).

         As used herein, the twenty conventional amino acids and their abbreviations follow conventional usage. *See* Golub *et al.* (eds.), Immunology - A Synthesis 2nd Ed., Sinauer

Associates (1991). Stereoisomers (*e.g.*, D-amino acids) of the twenty conventional amino acids, unnatural amino acids such as α-, α-disubstituted amino acids, N-alkyl amino acids, and other unconventional amino acids may also be suitable components for polypeptides of the present invention. Examples of unconventional amino acids include:

5   4-hydroxyproline, γ-carboxyglutamate, ε-N,N,N-trimethyllysine, ε-N-acetyllysine, O-phosphoserine, N-acetylserine, N-formylmethionine, 3-methylhistidine, 5-hydroxylysine, s-N-methylarginine, and other similar amino acids and imino acids (*e.g.*, 4-hydroxyproline). In the polypeptide notation used herein, the lefthand direction is the amino terminal direction and the right hand direction is the carboxy-terminal direction, in

10   accordance with standard usage and convention.

By "homology" or "homologous" when referring to a polypeptide of the present invention it is meant polypeptides from different organisms with a similar sequence to the encoded amino acid sequence of a OSP and a similar biological activity or function. Although two polypeptides are said to be "homologous," this does not imply that there is

15   necessarily an evolutionary relationship between the polypeptides. Instead, the term "homologous" is defined to mean that the two polypeptides have similar amino acid sequences and similar biological activities or functions. In a preferred embodiment, a homologous polypeptide is one that exhibits 50% sequence similarity to OSP, preferred is 60% sequence similarity, more preferred is 70% sequence similarity. Even more preferred

20   are homologous polypeptides that exhibit 80%, 85% or 90% sequence similarity to a OSP. In a yet more preferred embodiment, a homologous polypeptide exhibits 95%, 97%, 98% or 99% sequence similarity.

When "sequence similarity" is used in reference to polypeptides, it is recognized that residue positions that are not identical often differ by conservative amino acid

25   substitutions. In a preferred embodiment, a polypeptide that has "sequence similarity" comprises conservative or moderately conservative amino acid substitutions. A "conservative amino acid substitution" is one in which an amino acid residue is substituted by another amino acid residue having a side chain (R group) with similar chemical properties (*e.g.*, charge or hydrophobicity). In general, a conservative amino acid

30   substitution will not substantially change the functional properties of a protein. In cases where two or more amino acid sequences differ from each other by conservative substitutions, the percent sequence identity or degree of similarity may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this

adjustment are well known to those of skill in the art. *See, e.g.,* Pearson, *Methods Mol.*

Biol. 24: 307-31 (1994).

For instance, the following six groups each contain amino acids that are conservative substitutions for one another:

5   1)   Serine (S), Threonine (T);

              2)   Aspartic Acid (D), Glutamic Acid (E);

              3)   Asparagine (N), Glutamine (Q);

              4)   Arginine (R), Lysine (K);

              5)   Isoleucine (I), Leucine (L), Methionine (M), Alanine (A), Valine (V), and

10   6)   Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

Alternatively, a conservative replacement is any change having a positive value in the PAM250 log-likelihood matrix disclosed in Gonnet *et al.*, *Science* 256: 1443-45 (1992). A "moderately conservative" replacement is any change having a nonnegative value in the PAM250 log-likelihood matrix.

15   Sequence similarity for polypeptides, which is also referred to as sequence identity, is typically measured using sequence analysis software. Protein analysis software matches similar sequences using measures of similarity assigned to various substitutions, deletions and other modifications, including conservative amino acid substitutions. For instance, GCG contains programs such as "Gap" and "Bestfit" which can be used with

20   default parameters to determine sequence homology or sequence identity between closely related polypeptides, such as homologous polypeptides from different species of organisms or between a wild type protein and a mutein thereof. *See, e.g.,* GCG Version 6.1. Other programs include FASTA, discussed *supra.*

A preferred algorithm when comparing a sequence of the invention to a database

25   containing a large number of sequences from different organisms is the computer program BLAST, especially blastp or tblastn. *See, e.g.,* Altschul *et al.*, *J. Mol. Biol.* 215: 403-410 (1990); Altschul *et al.*, *Nucleic Acids Res.* 25:3389-402 (1997). Preferred parameters for blastp are:

              Expectation value: 10 (default)

30   Filter: seg (default)

              Cost to open a gap: 11 (default)

              Cost to extend a gap: 1 (default

              Max. alignments: 100 (default)

Word size:	11 (default)
No. of descriptions:	100 (default)
Penalty Matrix:	BLOSUM62

The length of polypeptide sequences compared for homology will generally be at

5   least about 16 amino acid residues, usually at least about 20 residues, more usually at least about 24 residues, typically at least about 28 residues, and preferably more than about 35 residues. When searching a database containing sequences from a large number of different organisms, it is preferable to compare amino acid sequences.

Algorithms other than blastp for database searching using amino acid sequences

10   are known in the art. For instance, polypeptide sequences can be compared using FASTA, a program in GCG Version 6.1. FASTA (*e.g.*, FASTA2 and FASTA3) provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences (Pearson (1990), *supra*; Pearson (2000), *supra*. For example, percent sequence identity between amino acid sequences can be determined using FASTA

15   with its default or recommended parameters (a word size of 2 and the PAM250 scoring matrix), as provided in GCG Version 6.1.

An "antibody" refers to an intact immunoglobulin, or to an antigen-binding portion thereof that competes with the intact antibody for specific binding to a molecular species, *e.g.*, a polypeptide of the instant invention. Antigen-binding portions may be produced by

20   recombinant DNA techniques or by enzymatic or chemical cleavage of intact antibodies. Antigen-binding portions include, *inter alia*, Fab, Fab', F(ab')$_2$, Fv, dAb, and complementarity determining region (CDR) fragments, single-chain antibodies (scFv), chimeric antibodies, diabodies and polypeptides that contain at least a portion of an immunoglobulin that is sufficient to confer specific antigen binding to the polypeptide. A

25   Fab fragment is a monovalent fragment consisting of the VL, VH, CL and CH1 domains; a F(ab')$_2$ fragment is a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; a Fd fragment consists of the VH and CH1 domains; a Fv fragment consists of the VL and VH domains of a single arm of an antibody; and a dAb fragment consists of a VH domain. *See, e.g.*, Ward *et al.*, *Nature* 341: 544-546 (1989).

30   By "bind specifically" and "specific binding" as used herein it is meant the ability of the antibody to bind to a first molecular species in preference to binding to other molecular species with which the antibody and first molecular species are admixed. An

antibody is said specifically to "recognize" a first molecular species when it can bind specifically to that first molecular species.

A single-chain antibody (scFv) is an antibody in which VL and VH regions are paired to form a monovalent molecule via a synthetic linker that enables them to be made as a single protein chain. *See, e.g.*, Bird *et al.*, *Science* 242: 423-426 (1988); Huston *et al.*, *Proc. Natl. Acad. Sci. USA* 85: 5879-5883 (1988). Diabodies are bivalent, bispecific antibodies in which VH and VL domains are expressed on a single polypeptide chain, but using a linker that is too short to allow for pairing between the two domains on the same chain, thereby forcing the domains to pair with complementary domains of another chain and creating two antigen binding sites. *See e.g.*, Holliger *et al.*, *Proc. Natl. Acad. Sci. USA* 90: 6444-6448 (1993); Poljak *et al.*, *Structure* 2: 1121-1123 (1994). One or more CDRs may be incorporated into a molecule either covalently or noncovalently to make it an immunoadhesin. An immunoadhesin may incorporate the CDR(s) as part of a larger polypeptide chain, may covalently link the CDR(s) to another polypeptide chain, or may incorporate the CDR(s) noncovalently. The CDRs permit the immunoadhesin to specifically bind to a particular antigen of interest. A chimeric antibody is an antibody that contains one or more regions from one antibody and one or more regions from one or more other antibodies.

An antibody may have one or more binding sites. If there is more than one binding site, the binding sites may be identical to one another or may be different. For instance, a naturally occurring immunoglobulin has two identical binding sites, a single-chain antibody or Fab fragment has one binding site, while a "bispecific" or "bifunctional" antibody has two different binding sites.

An "isolated antibody" is an antibody that (1) is not associated with naturally-associated components, including other naturally-associated antibodies, that accompany it in its native state, (2) is free of other proteins from the same species, (3) is expressed by a cell from a different species, or (4) does not occur in nature. It is known that purified proteins, including purified antibodies, may be stabilized with non-naturally-associated components. The non-naturally-associated component may be a protein, such as albumin (*e.g.*, BSA) or a chemical such as polyethylene glycol (PEG).

A "neutralizing antibody" or "an inhibitory antibody" is an antibody that inhibits the activity of a polypeptide or blocks the binding of a polypeptide to a ligand that

normally binds to it. An "activating antibody" is an antibody that increases the activity of a polypeptide.

The term "epitope" includes any protein determinant capable of specific binding to an immunoglobulin or T-cell receptor. Epitopic determinants usually consist of

5   chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three-dimensional structural characteristics, as well as specific charge characteristics. An antibody is said to specifically bind an antigen when the dissociation constant is less than1 µM, preferably less than100 nM and most preferably less than 10 nM.

10   The term "patient" includes human and veterinary subjects.

Throughout this specification and claims, the word "comprise," or variations such as "comprises" or "comprising," will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

The term "ovarian specific" refers to a nucleic acid molecule or polypeptide that is

15   expressed predominantly in the ovarian as compared to other tissues in the body. In a preferred embodiment, a "ovarian specific" nucleic acid molecule or polypeptide is detected at a level that is 1.5-fold higher than any other tissue in the body. In a more preferred embodiment, the "ovarian specific" nucleic acid molecule or polypeptide is detected at a level that is 2-fold higher than any other tissue in the body, more preferably

20   5-fold higher, still more preferably at least 10-fold, 15-fold, 20-fold, 25-fold, 50-fold or 100-fold higher than any other tissue in the body. Nucleic acid molecule levels may be measured by nucleic acid hybridization, such as Northern blot hybridization, or quantitative PCR. Polypeptide levels may be measured by any method known to accurately quantitate protein levels, such as Western blot analysis.

25   **Nucleic Acid Molecules. Regulatory Sequences, Vectors, Host Cells and Recombinant Methods of Making Polypeptides**

Nucleic Acid Molecules

One aspect of the invention provides isolated nucleic acid molecules that are specific to the ovarian or to ovarian cells or tissue or that are derived from such nucleic

30   acid molecules. These isolated ovarian specific nucleic acids (OSNAs) may comprise cDNA genomic DNA, RNA, or a combination thereof, a fragment of one of these nucleic acids, or may be a non-naturally occurring nucleic acid molecule. A OSNA may be

derived from an animal. In a preferred embodiment, the OSNA is derived from a human or other mammal. In a more preferred embodiment, the OSNA is derived from a human or other primate. In an even more preferred embodiment, the OSNA is derived from a human.

5   In a preferred embodiment, the nucleic acid molecule encodes a polypeptide that is specific to ovarian, a ovarian-specific polypeptide (OSP). In a more preferred embodiment, the nucleic acid molecule encodes a polypeptide that comprises an amino acid sequence of SEQ ID NO: 100-192. In another highly preferred embodiment, the nucleic acid molecule comprises a nucleic acid sequence of SEQ ID NO: 1-99. Nucleotide

10   sequences of the instantly-described nucleic acid molecules were determined by assembling several DNA molecules from either public or proprietary databases. Some of the underlying DNA sequences are the result, directly or indirectly, of at least one enzymatic polymerization reaction (*e.g.*, reverse transcription and/or polymerase chain reaction) using an automated sequencer (such as the MegaBACE™ 1000, Amersham

15   Biosciences, Sunnyvale, CA, USA).

         Nucleic acid molecules of the present invention may also comprise sequences that selectively hybridizes to a nucleic acid molecule encoding a OSNA or a complement or antisense thereof. The hybridizing nucleic acid molecule may or may not encode a polypeptide or may or may not encode a OSP. However, in a preferred embodiment, the

20   hybridizing nucleic acid molecule encodes a OSP. In a more preferred embodiment, the invention provides a nucleic acid molecule that selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule that encodes a polypeptide comprising an amino acid sequence of SEQ ID NO: 100-192. In an even more preferred embodiment, the invention provides a nucleic acid molecule that selectively hybridizes to

25   a nucleic acid molecule comprising the nucleic acid sequence of SEQ ID NO: 1-99 or the antisense sequence thereof. Preferably, the nucleic acid molecule selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a OSP under low stringency conditions. More preferably, the nucleic acid molecule selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic

30   acid molecule encoding a OSP under moderate stringency conditions. Most preferably, the nucleic acid molecule selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a OSP under high stringency conditions. In a preferred embodiment, the nucleic acid molecule hybridizes under low, moderate or high

stringency conditions to a nucleic acid molecule or the antisense sequence of a nucleic

acid molecule encoding a polypeptide comprising an amino acid sequence of SEQ ID NO:

100-192. In a more preferred embodiment, the nucleic acid molecule hybridizes under

low, moderate or high stringency conditions to a nucleic acid molecule or the antisense

5   sequence of a nucleic acid molecule comprising a nucleic acid sequence selected from

SEQ ID NO: 1-99.

Nucleic acid molecules of the present invention may also comprise nucleic acid

sequences that exhibit substantial sequence similarity to a nucleic acid encoding a OSP or

a complement of the encoding nucleic acid molecule. In this embodiment, it is preferred

10   that the nucleic acid molecule exhibit substantial sequence similarity to a nucleic acid

molecule encoding human OSP. More preferred is a nucleic acid molecule exhibiting

substantial sequence similarity to a nucleic acid molecule encoding a polypeptide having

an amino acid sequence of SEQ ID NO: 100-192. By substantial sequence similarity it is

meant a nucleic acid molecule having at least 60% sequence identity with a nucleic acid

15   molecule encoding a OSP, such as a polypeptide having an amino acid sequence of SEQ

ID NO: 100-192, more preferably at least 70%, even more preferably at least 80% and

even more preferably at least 85%. In a more preferred embodiment, the similar nucleic

acid molecule is one that has at least 90% sequence identity with a nucleic acid molecule

encoding a OSP, more preferably at least 95%, more preferably at least 97%, even more

20   preferably at least 98%, and still more preferably at least 99%. Most preferred in this

embodiment is a nucleic acid molecule that has at least 99.5%, 99.6%, 99.7%, 99.8% or

99.9% sequence identity with a nucleic acid molecule encoding a OSP.

The nucleic acid molecules of the present invention are also inclusive of those

exhibiting substantial sequence similarity to a OSNA or its complement. In this

25   embodiment, it is preferred that the nucleic acid molecule exhibit substantial sequence

similarity to a nucleic acid molecule having a nucleic acid sequence of SEQ ID NO: 1-99.

By substantial sequence similarity it is meant a nucleic acid molecule that has at least 60%

sequence identity with a OSNA, such as one having a nucleic acid sequence of SEQ ID

NO: 1-99, more preferably at least 70%, even more preferably at least 80% and even more

30   preferably at least 85%. More preferred is a nucleic acid molecule that has at least 90%

sequence identity with a OSNA, more preferably at least 95%, more preferably at least

97%, even more preferably at least 98%, and still more preferably at least 99%. Most

preferred is a nucleic acid molecule that has at least 99.5%, 99.6%, 99.7%, 99.8% or 99.9% sequence identity with a OSNA.

Nucleic acid molecules that exhibit substantial sequence similarity are inclusive of sequences that exhibit sequence identity over their entire length to a OSNA or to a nucleic acid molecule encoding a OSP, as well as sequences that are similar over only a part of its length. In this case, the part is at least 50 nucleotides of the OSNA or the nucleic acid molecule encoding a OSP, preferably at least 100 nucleotides, more preferably at least 150 or 200 nucleotides, even more preferably at least 250 or 300 nucleotides, still more preferably at least 400 or 500 nucleotides.

The substantially similar nucleic acid molecule may be a naturally occurring one that is derived from another species, especially one derived from another primate, wherein the similar nucleic acid molecule encodes an amino acid sequence that exhibits significant sequence identity to that of SEQ ID NO: 100-192 or demonstrates significant sequence identity to the nucleotide sequence of SEQ ID NO: 1-99. The similar nucleic acid molecule may also be a naturally occurring nucleic acid molecule from a human, when the OSNA is a member of a gene family. The similar nucleic acid molecule may also be a naturally occurring nucleic acid molecule derived from a non-primate, mammalian species, including without limitation, domesticated species, *e.g.*, dog, cat, mouse, rat, rabbit, hamster, cow, horse and pig; and wild animals, *e.g.*, monkey, fox, lions, tigers, bears, giraffes, zebras, etc. The substantially similar nucleic acid molecule may also be a naturally occurring nucleic acid molecule derived from a non-mammalian species, such as birds or reptiles. The naturally occurring substantially similar nucleic acid molecule may be isolated directly from humans or other species. In another embodiment, the substantially similar nucleic acid molecule may be one that is experimentally produced by random mutation of a nucleic acid molecule. In another embodiment, the substantially similar nucleic acid molecule may be one that is experimentally produced by directed mutation of a OSNA. In a preferred embodiment, the substantially similar nucleic acid molecule is an OSNA.

The nucleic acid molecules of the present invention are also inclusive of allelic variants of a OSNA or a nucleic acid encoding a OSP. For example, single nucleotide polymorphisms (SNPs) occur frequently in eukaryotic genomes and the sequence determined from one individual of a species may differ from other allelic forms present within the population. More than 1.4 million SNPs have already identified in the human

genome, International Human Genome Sequencing Consortium, *Nature* 409: 860-921

(2001) — Variants with small deletions and insertions of more than a single nucleotide are

also found in the general population, and often do not alter the function of the protein. In

addition, amino acid substitutions occur frequently among natural allelic variants, and

5   often do not substantially change protein function.

In a preferred embodiment, the allelic variant is a variant of a gene, wherein the

gene is transcribed into an mRNA that encodes a OSP. In a more preferred embodiment,

the gene is transcribed into an mRNA that encodes a OSP comprising an amino acid

sequence of SEQ ID NO: 100-192. In another preferred embodiment, the allelic variant is

10   a variant of a gene, wherein the gene is transcribed into an mRNA that is a OSNA. In a

more preferred embodiment, the gene is transcribed into an mRNA that comprises the

nucleic acid sequence of SEQ ID NO: 1-99. Also preferred is that the allelic variant is a

naturally occurring allelic variant in the species of interest, particularly human.

Nucleic acid molecules of the present invention are also inclusive of nucleic acid

15   sequences comprising a part of a nucleic acid sequence of the instant invention. The part

may or may not encode a polypeptide, and may or may not encode a polypeptide that is a

OSP. In a preferred embodiment, the part encodes a OSP. In one embodiment, the

nucleic acid molecule comprises a part of a OSNA. In another embodiment, the nucleic

acid molecule comprises a part of a nucleic acid molecule that hybridizes or exhibits

20   substantial sequence similarity to a OSNA. In another embodiment, the nucleic acid

molecule comprises a part of a nucleic acid molecule that is an allelic variant of a OSNA.

In yet another embodiment, the nucleic acid molecule comprises a part of a nucleic acid

molecule that encodes a OSP. A part comprises at least 10 nucleotides, more preferably at

least 15, 17, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400 or

25   500 nucleotides. The maximum size of a nucleic acid part is one nucleotide shorter than

the sequence of the nucleic acid molecule encoding the full-length protein.

Nucleic acid molecules of the present invention are also inclusive of nucleic acid

sequences that encode fusion proteins, homologous proteins, polypeptide fragments,

muteins and polypeptide analogs, as described *infra*.

30   Nucleic acid molecules of the present invention are also inclusive of nucleic acid

sequences containing modifications of the native nucleic acid molecule. Examples of such

modifications include, but are not limited to, nonnative internucleoside bonds, post-

synthetic modifications or altered nucleotide analogues. One having ordinary skill in the

art would recognize that the type of modification that may be made will depend upon the intended use of the nucleic acid molecule. For instance, when the nucleic acid molecule is used as a hybridization probe, the range of such modifications will be limited to those that permit sequence-discriminating base pairing of the resulting nucleic acid. When used to

5   direct expression of RNA or protein *in vitro* or *in vivo*, the range of such modifications will be limited to those that permit the nucleic acid to function properly as a polymerization substrate. When the isolated nucleic acid is used as a therapeutic agent, the modifications will be limited to those that do not confer toxicity upon the isolated nucleic acid.

10   Accordingly, in one embodiment, a nucleic acid molecule may include nucleotide analogues that incorporate labels that are directly detectable, such as radiolabels or fluorophores, or nucleotide analogues that incorporate labels that can be visualized in a subsequent reaction, such as biotin or various haptens. The labeled nucleic acid molecules are particularly useful as hybridization probes.

15   Common radiolabeled analogues include those labeled with ^{33}P, ^{32}P, and ^{35}S, such as $\alpha\text{-}^{32}P\text{-dATP}$, $\alpha\text{-}^{32}P\text{-dCTP}$, $\alpha\text{-}^{32}P\text{-dGTP}$, $\alpha\text{-}^{32}P\text{-dTTP}$, $\alpha\text{-}^{32}P\text{-3'dATP}$, $\alpha\text{-}^{32}P\text{-ATP}$, $\alpha\text{-}^{32}P\text{-CTP}$, $\alpha\text{-}^{32}P\text{-GTP}$, $\alpha\text{-}^{32}P\text{-UTP}$, $\alpha\text{-}^{35}S\text{-dATP}$, $\gamma\text{-}^{35}S\text{-GTP}$, $\gamma\text{-}^{33}P\text{-dATP}$, and the like.

Commercially available fluorescent nucleotide analogues readily incorporated into the nucleic acids of the present invention include Cy3-dCTP, Cy3-dUTP, Cy5-dCTP, Cy3-

20   dUTP (Amersham Biosciences, Piscataway, New Jersey, USA), fluorescein-12-dUTP, tetramethylrhodamine-6-dUTP, Texas Red®-5-dUTP, Cascade Blue®-7-dUTP, BODIPY® FL-14-dUTP, BODIPY® TMR-14-dUTP, BODIPY® TR-14-dUTP, Rhodamine Green™-5-dUTP, Oregon Green® 488-5-dUTP, Texas Red®-12-dUTP, BODIPY® 630/650-14-dUTP, BODIPY® 650/665-14-dUTP, Alexa Fluor® 488-5-dUTP,

25   Alexa Fluor® 532-5-dUTP, Alexa Fluor® 568-5-dUTP, Alexa Fluor® 594-5-dUTP, Alexa Fluor® 546-14-dUTP, fluorescein-12-UTP, tetramethylrhodamine-6-UTP, Texas Red®-5-UTP, Cascade Blue®-7-UTP, BODIPY® FL-14-UTP, BODIPY® TMR-14-UTP, BODIPY® TR-14-UTP, Rhodamine Green™-5-UTP, Alexa Fluor® 488-5-UTP, Alexa Fluor® 546-14-UTP (Molecular Probes, Inc. Eugene, OR, USA). One may also custom

30   synthesize nucleotides having other fluorophores. *See* Henegariu *et al.*, *Nature Biotechnol.* 18: 345-348 (2000).

Haptens that are commonly conjugated to nucleotides for subsequent labeling include biotin (biotin-11-dUTP, Molecular Probes, Inc., Eugene, OR, USA;

biotin-21-UTP, biotin-21-dUTP, Clontech Laboratories, Inc., Palo Alto, CA, USA),

digoxigenin (DIG-11-dUTP, alkali labile, DIG-11-UTP, Roche Diagnostics Corp.,

Indianapolis, IN, USA), and dinitrophenyl (dinitrophenyl-11-dUTP, Molecular Probes,

Inc., Eugene, OR, USA).

5   Nucleic acid molecules of the present invention can be labeled by incorporation of

labeled nucleotide analogues into the nucleic acid. Such analogues can be incorporated by

enzymatic polymerization, such as by nick translation, random priming, polymerase chain

reaction (PCR), terminal transferase tailing, and end-filling of overhangs, for DNA

molecules, and *in vitro* transcription driven, *e.g.*, from phage promoters, such as T7, T3,

10   and SP6, for RNA molecules. Commercial kits are readily available for each such

labeling approach. Analogues can also be incorporated during automated solid phase

chemical synthesis. Labels can also be incorporated after nucleic acid synthesis, with the

5' phosphate and 3' hydroxyl providing convenient sites for post-synthetic covalent

attachment of detectable labels.

15   Other post-synthetic approaches also permit internal labeling of nucleic acids. For

example, fluorophores can be attached using a cisplatin reagent that reacts with the N7 of

guanine residues (and, to a lesser extent, adenine bases) in DNA, RNA, and Peptide

Nucleic Acids (PNA) to provide a stable coordination complex between the nucleic acid

and fluorophore label (Universal Linkage System) (available from Molecular Probes, Inc.,

20   Eugene, OR, USA and Amersham Pharmacia Biotech, Piscataway, NJ, USA); *see* Alers *et*

al., *Genes, Chromosomes & Cancer* 25: 301- 305 (1999); Jelsma *et al.*, *J. NIH Res.* 5: 82

(1994); Van Belkum *et al.*, *BioTechniques* 16: 148-153 (1994). Alternatively, nucleic

acids can be labeled using a disulfide-containing linker (FastTag™ Reagent, Vector

Laboratories, Inc., Burlingame, CA, USA) that is photo- or thermally coupled to the target

25   nucleic acid using aryl azide chemistry; after reduction, a free thiol is available for

coupling to a hapten, fluorophore, sugar, affinity ligand, or other marker.

One or more independent or interacting labels can be incorporated into the nucleic

acid molecules of the present invention. For example, both a fluorophore and a moiety

that in proximity thereto acts to quench fluorescence can be included to report specific

30   hybridization through release of fluorescence quenching or to report exonucleotidic

excision. *See, e.g.,* Tyagi *et al.*, *Nature Biotechnol.* 14: 303-308 (1996); Tyagi *et al.*,

Nature Biotechnol. 16: 49-53 (1998); Sokol *et al.*, *Proc. Natl. Acad. Sci. USA* 95:

11538-11543 (1998); Kostrikis *et al.*, *Science* 279: 1228-1229 (1998); Marras *et al.*,

38

Genet. Anal. 14: 151-156 (1999); Holland *et al., Proc. Natl. Acad. Sci. USA* 88: 7276-7280 (1991); Heid *et al., Genome Res.* 6(10): 986-94 (1996); Kuimelis *et al., Nucleic Acids Symp. Ser.* (37): 255-6 (1997); and U.S. Patent Nos. 5,846,726, 5,925,517, 5,925,517, 5,723,591 and 5,538,848, the disclosures of which are incorporated herein by

5   reference in their entireties.

Nucleic acid molecules of the present invention may also be modified by altering one or more native phosphodiester internucleoside bonds to more nuclease-resistant, internucleoside bonds. *See* Hartmann *et al.* (eds.), <u>Manual of Antisense Methodology: Perspectives in Antisense Science</u>, Kluwer Law International (1999); Stein *et al.* (eds.),

10   <u>Applied Antisense Oligonucleotide Technology</u>, Wiley-Liss (1998); Chadwick *et al.* (eds.), <u>Oligonucleotides as Therapeutic Agents – Symposium No. 209</u>, John Wiley & Son Ltd (1997). Such altered internucleoside bonds are often desired for techniques or for targeted gene correction, Gamper *et al., Nucl. Acids Res.* 28(21): 4332-4339 (2000). For double stranded RNA inhibition which may utilize either natural ds RNA or ds RNA

15   modified in its, sugar, phosphate or base, *see* Hannon, *Nature* 418(11): 244-251 (2002); Fire *et al.* in WO 99/32619; Tuschl *et al.* in US2002/0086356; Kruetzer *et al.* in WO 00/44895, the disclosures of which are incorporated herein by reference in their entirety;. For circular antisense, see Kool in U.S. Patent No. 5,426,180, the disclosure of which is incorporated herein by reference in its entirety.

20   Modified oligonucleotide backbones include, without limitation, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates,

25   thionoalkylphosphonates, thionoalkylphosphotriesters, and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein the adjacent pairs of nucleoside units are linked 3'-5' to 5'-3' or 2'-5' to 5'-2'. Representative U.S. Patents that teach the preparation of the above phosphorus-containing linkages include, but are not limited to, U.S. Patent Nos. 3,687,808; 4,469,863; 4,476,301;

30   5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; and 5,625,050, the disclosures of

which are incorporated herein by reference in their entireties. In a preferred embodiment, the modified internucleoside linkages may be used for antisense techniques.

Other modified oligonucleotide backbones do not include a phosphorus atom, but have backbones that are formed by short chain alkyl or cycloalkyl internucleoside

5   linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; alkene

10   containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH_2 component parts. Representative U.S. patents that teach the preparation of the above backbones include, but are not limited to, U.S. Patent Nos. 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564;

15   5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437 and 5,677,439; the disclosures of which are incorporated herein by reference in their entireties.

In other preferred nucleic acid molecules, both the sugar and the internucleoside

20   linkage are replaced with novel groups, such as peptide nucleic acids (PNA). In PNA compounds, the phosphodiester backbone of the nucleic acid is replaced with an amide-containing backbone, in particular by repeating N-(2-aminoethyl) glycine units linked by amide bonds. Nucleobases are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone, typically by methylene carbonyl linkages. PNA can be

25   synthesized using a modified peptide synthesis protocol. PNA oligomers can be synthesized by both Fmoc and tBoc methods. Representative U.S. patents that teach the preparation of PNA compounds include, but are not limited to, U.S. Patent Nos. 5,539,082; 5,714,331; and 5,719,262, each of which is herein incorporated by reference in its entirety. Automated PNA synthesis is readily achievable on commercial synthesizers

30   (see, e.g., "PNA User's Guide," Rev. 2, February 1998, Perseptive Biosystems Part No. 60138, Applied Biosystems, Inc., Foster City, CA). PNA molecules are advantageous for a number of reasons. First, because the PNA backbone is uncharged, PNA/DNA and PNA/RNA duplexes have a higher thermal stability than is found in DNA/DNA and

DNA/RNA duplexes. The Tm of a PNA/DNA or PNA/RNA duplex is generally 1°C higher per base pair than the Tm of the corresponding DNA/DNA or DNA/RNA duplex (in 100 mM NaCl). Second, PNA molecules can also form stable PNA/DNA complexes at low ionic strength, under conditions in which DNA/DNA duplex formation does not

5   occur. Third, PNA also demonstrates greater specificity in binding to complementary DNA because a PNA/DNA mismatch is more destabilizing than DNA/DNA mismatch. A single mismatch in mixed a PNA/DNA 15-mer lowers the Tm by 8–20°C (15°C on average). In the corresponding DNA/DNA duplexes, a single mismatch lowers the Tm by 4–16°C (11°C on average). Because PNA probes can be significantly shorter than DNA

10 probes, their specificity is greater. Fourth, PNA oligomers are resistant to degradation by enzymes, and the lifetime of these compounds is extended both *in vivo* and *in vitro* because nucleases and proteases do not recognize the PNA polyamide backbone with nucleobase sidechains. *See, e.g.,* Ray *et al., FASEB J.* 14(9): 1041-60 (2000); Nielsen *et al., Pharmacol Toxicol.* 86(1): 3-7 (2000); Larsen *et al., Biochim Biophys Acta.* 1489(1):

15 159-66 (1999); Nielsen, *Curr. Opin. Struct. Biol.* 9(3): 353-7 (1999), and Nielsen, *Curr. Opin. Biotechnol.* 10(1): 71-5 (1999).

Nucleic acid molecules may be modified compared to their native structure throughout the length of the nucleic acid molecule or can be localized to discrete portions thereof. As an example of the latter, chimeric nucleic acids can be synthesized that have

20 discrete DNA and RNA domains and that can be used for targeted gene repair and modified PCR reactions, as further described in, Misra *et al., Biochem.* 37: 1917-1925 (1998); and Finn *et al., Nucl. Acids Res.* 24: 3357-3363 (1996), and U.S. Patent Nos. 5,760,012 and 5,731,181, the disclosures of which are incorporated herein by reference in their entireties.

25 Unless otherwise specified, nucleic acid molecules of the present invention can include any topological conformation appropriate to the desired use; the term thus explicitly comprehends, among others, single-stranded, double-stranded, triplexed, quadruplexed, partially double-stranded, partially-triplexed, partially-quadruplexed, branched, hairpinned, circular, and padlocked conformations. Padlock conformations and

30 their utilities are further described in Banér *et al., Curr. Opin. Biotechnol.* 12: 11-15 (2001); Escude *et al., Proc. Natl. Acad. Sci. USA* 14: 96(19):10603-7 (1999); and Nilsson *et al., Science* 265(5181): 2085-8 (1994). Triplex and quadruplex conformations, and their utilities, are reviewed in Praseuth *et al., Biochim. Biophys. Acta.* 1489(1): 181-206

(1999); Fox, *Curr. Med. Chem.* 7(1): 17-37 (2000); Kochetkova *et al.*, *Methods Mol. Biol.* 130: 189-201 (2000); Chan *et al.*, *J. Mol. Med.* 75(4): 267-82 (1997); Rowley *et al.*, *Mol Med* 5(10): 693-700 (1999); Kool, *Annu Rev Biophys Biomol Struct.* 25: 1-28 (1996).

5   *Methods for Using Nucleic Acid Molecules as Probes and Primers*

The isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize, and quantify hybridizing nucleic acids in, and isolate hybridizing nucleic acids from, both genomic and transcript-derived nucleic acid samples. When free in solution, such probes are typically, but not invariably, detectably

10   labeled; bound to a substrate, as in a microarray, such probes are typically, but not invariably unlabeled.

In one embodiment, the isolated nucleic acid molecules of the present invention can be used as probes to detect and characterize gross alterations in the gene of a OSNA, such as deletions, insertions, translocations, and duplications of the OSNA genomic locus

15   through fluorescence *in situ* hybridization (FISH) to chromosome spreads. *See, e.g.,* Andreeff *et al.* (eds.), Introduction to Fluorescence *In Situ* Hybridization: Principles and Clinical Applications, John Wiley & Sons (1999). The isolated nucleic acid molecules of the present invention can be used as probes to assess smaller genomic alterations using, *e.g.,* Southern blot detection of restriction fragment length polymorphisms. The isolated

20   nucleic acid molecules of the present invention can be used as probes to isolate genomic clones that include a nucleic acid molecule of the present invention, which thereafter can be restriction mapped and sequenced to identify deletions, insertions, translocations, and substitutions (single nucleotide polymorphisms, SNPs) at the sequence level.

Alternatively, detection techniques such as molecular beacons may be used, see Kostrikis

25   *et al. Science* 279:1228-1229 (1998).

The isolated nucleic acid molecules of the present invention can be also be used as probes to detect, characterize, and quantify OSNA in, and isolate OSNA from, transcript-derived nucleic acid samples. In one embodiment, the isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize by length,

30   and quantify mRNA by Northern blot of total or poly-A$^+$- selected RNA samples. In another embodiment, the isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize by location, and quantify mRNA by *in situ* hybridization to tissue sections. *See, e.g.,* Schwarchzacher *et al.*, In Situ

Hybridization, Springer-Verlag New York (2000). In another preferred embodiment, the isolated nucleic acid molecules of the present invention can be used as hybridization probes to measure the representation of clones in a cDNA library or to isolate hybridizing nucleic acid molecules acids from cDNA libraries, permitting sequence level

5   characterization of mRNAs that hybridize to OSNAs, including, without limitations, identification of deletions, insertions, substitutions, truncations, alternatively spliced forms and single nucleotide polymorphisms. In yet another preferred embodiment, the nucleic acid molecules of the instant invention may be used in microarrays.

All of the aforementioned probe techniques are well within the skill in the art, and

10   are described at greater length in standard texts such as Sambrook (2001), *supra*; Ausubel (1999), *supra*; and Walker *et al.* (eds.), The Nucleic Acids Protocols Handbook, Humana Press (2000).

In another embodiment, a nucleic acid molecule of the invention may be used as a probe or primer to identify and/or amplify a second nucleic acid molecule that selectively

15   hybridizes to the nucleic acid molecule of the invention. In this embodiment, it is preferred that the probe or primer be derived from a nucleic acid molecule encoding a OSP. More preferably, the probe or primer is derived from a nucleic acid molecule encoding a polypeptide having an amino acid sequence of SEQ ID NO: 100-192. Also preferred are probes or primers derived from a OSNA. More preferred are probes or

20   primers derived from a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-99.

In general, a probe or primer is at least 10 nucleotides in length, more preferably at least 12, more preferably at least 14 and even more preferably at least 16 or 17 nucleotides in length. In an even more preferred embodiment, the probe or primer is at least 18

25   nucleotides in length, even more preferably at least 20 nucleotides and even more preferably at least 22 nucleotides in length. Primers and probes may also be longer in length. For instance, a probe or primer may be 25 nucleotides in length, or may be 30, 40 or 50 nucleotides in length. Methods of performing nucleic acid hybridization using oligonucleotide probes are well known in the art. *See, e.g.*, Sambrook *et al.*, 1989, *supra*,

30   Chapter 11 and pp. 11.31-11.32 and 11.40-11.44, which describes radiolabeling of short probes, and pp. 11.45-11.53, which describe hybridization conditions for oligonucleotide probes, including specific conditions for probe hybridization (pp. 11.50-11.51).

Methods of performing primer-directed amplification are also well known in the
art. Methods for performing the polymerase chain reaction (PCR) are compiled, *inter alia*,
in McPherson, PCR Basics: From Background to Bench, Springer Verlag (2000); Innis *et
al.* (eds.), PCR Applications: Protocols for Functional Genomics, Academic Press (1999);

5   Gelfand *et al.* (eds.), PCR Strategies, Academic Press (1998); Newton *et al.*, PCR,
Springer-Verlag New York (1997); Burke (ed.), PCR: Essential Techniques, John Wiley
& Son Ltd (1996); White (ed.), PCR Cloning Protocols: From Molecular Cloning to
Genetic Engineering, Vol. 67, Humana Press (1996); and McPherson *et al.* (eds.), PCR 2:
A Practical Approach, Oxford University Press, Inc. (1995). Methods for performing RT-

10   PCR are collected, *e.g.*, in Siebert *et al.* (eds.), Gene Cloning and Analysis by RT-PCR,
Eaton Publishing Company/Bio Techniques Books Division, 1998; and Siebert (ed.), PCR
Technique:RT-PCR, Eaton Publishing Company/ BioTechniques Books (1995).

PCR and hybridization methods may be used to identify and/or isolate nucleic acid
molecules of the present invention including allelic variants, homologous nucleic acid

15   molecules and fragments. PCR and hybridization methods may also be used to identify,
amplify and/or isolate nucleic acid molecules of the present invention that encode
homologous proteins, analogs, fusion protein or muteins of the invention. Nucleic acid
primers as described herein can be used to prime amplification of nucleic acid molecules
of the invention, using transcript-derived or genomic DNA as template.

20   These nucleic acid primers can also be used, for example, to prime single base
extension (SBE) for SNP detection (*See, e.g.*, U.S. Pat. No. 6,004,744, the disclosure of
which is incorporated herein by reference in its entirety).

Isothermal amplification approaches, such as rolling circle amplification, are also
now well-described. *See, e.g.*, Schweitzer *et al.*, *Curr. Opin. Biotechnol.* 12(1): 21-7

25   (2001); international patent publications WO 97/19193 and WO 00/15779, and U.S. Patent
Nos. 5,854,033 and 5,714,320, the disclosures of which are incorporated herein by
reference in their entireties. Rolling circle amplification can be combined with other
techniques to facilitate SNP detection. *See, e.g.*, Lizardi *et al.*, *Nature Genet.* 19(3):
225-32 (1998).

30   Nucleic acid molecules of the present invention may be bound to a substrate either
covalently or noncovalently. The substrate can be porous or solid, planar or non-planar,
unitary or distributed. The bound nucleic acid molecules may be used as hybridization

probes, and may be labeled or unlabeled. In a preferred embodiment, the bound nucleic acid molecules are unlabeled.

In one embodiment, the nucleic acid molecule of the present invention is bound to a porous substrate, e.g., a membrane, typically comprising nitrocellulose, nylon, or

5   positively charged derivatized nylon. The nucleic acid molecule of the present invention can be used to detect a hybridizing nucleic acid molecule that is present within a labeled nucleic acid sample, e.g., a sample of transcript-derived nucleic acids. In another embodiment, the nucleic acid molecule is bound to a solid substrate, including, without limitation, glass, amorphous silicon, crystalline silicon or plastics. Examples of plastics

10   include, without limitation, polymethylacrylic, polyethylene, polypropylene, polyacrylate, polymethylmethacrylate, polyvinylchloride, polytetrafluoroethylene, polystyrene, polycarbonate, polyacetal, polysulfone, celluloseacetate, cellulosenitrate, nitrocellulose, or mixtures thereof. The solid substrate may be any shape, including rectangular, disk-like and spherical. In a preferred embodiment, the solid substrate is a microscope slide or

15   slide-shaped substrate.

The nucleic acid molecule of the present invention can be attached covalently to a surface of the support substrate or applied to a derivatized surface in a chaotropic agent that facilitates denaturation and adherence by presumed noncovalent interactions, or some combination thereof. The nucleic acid molecule of the present invention can be bound to a

20   substrate to which a plurality of other nucleic acids are concurrently bound, hybridization to each of the plurality of bound nucleic acids being separately detectable. At low density, e.g. on a porous membrane, these substrate-bound collections are typically denominated macroarrays; at higher density, typically on a solid support, such as glass, these substrate bound collections of plural nucleic acids are colloquially termed microarrays. As used

25   herein, the term microarray includes arrays of all densities. It is, therefore, another aspect of the invention to provide microarrays that comprise one or more of the nucleic acid molecules of the present invention.

In yet another embodiment, the invention is directed to single exon probes based on the OSNAs disclosed herein.

30

Expression Vectors, Host Cells and Recombinant Methods of Producing
Polypeptides

Another aspect of the present invention provides vectors that comprise one or more of the isolated nucleic acid molecules of the present invention, and host cells in which such vectors have been introduced.

The vectors can be used, *inter alia*, for propagating the nucleic acid molecules of

5   the present invention in host cells (cloning vectors), for shuttling the nucleic acid molecules of the present invention between host cells derived from disparate organisms (shuttle vectors), for inserting the nucleic acid molecules of the present invention into host cell chromosomes (insertion vectors), for expressing sense or antisense RNA transcripts of the nucleic acid molecules of the present invention *in vitro* or within a host cell, and for

10   expressing polypeptides encoded by the nucleic acid molecules of the present invention, alone or as fusion proteins with heterologous polypeptides (expression vectors). Vectors are by now well known in the art, and are described, *inter alia*, in Jones *et al.* (eds.), Vectors: Cloning Applications: Essential Techniques (Essential Techniques Series), John Wiley & Son Ltd. (1998); Jones *et al.* (eds.), Vectors: Expression Systems: Essential

15   Techniques (Essential Techniques Series), John Wiley & Son Ltd. (1998); Gacesa *et al.*, Vectors: Essential Data, John Wiley & Sons Ltd. (1995); Cid-Arregui (eds.), Viral Vectors: Basic Science and Gene Therapy, Eaton Publishing Co. (2000); Sambrook (2001), *supra*; Ausubel (1999), *supra*. Furthermore, a variety of vectors are available commercially. Use of existing vectors and modifications thereof are well within the skill

20   in the art. Thus, only basic features need be described here.

Nucleic acid sequences may be expressed by operatively linking them to an expression control sequence in an appropriate expression vector and employing that expression vector to transform an appropriate unicellular host. Expression control sequences are sequences that control the transcription, post-transcriptional events and

25   translation of nucleic acid sequences. Such operative linking of a nucleic sequence of this invention to an expression control sequence, of course, includes, if not already part of the nucleic acid sequence, the provision of a translation initiation codon, ATG or GTG, in the correct reading frame upstream of the nucleic acid sequence.

A wide variety of host/expression vector combinations may be employed in

30   expressing the nucleic acid sequences of this invention. Useful expression vectors, for example, may consist of segments of chromosomal, non-chromosomal and synthetic nucleic acid sequences.

In one embodiment, prokaryotic cells may be used with an appropriate vector. Prokaryotic host cells are often used for cloning and expression. In a preferred embodiment, prokaryotic host cells include *E. coli, Pseudomonas, Bacillus* and *Streptomyces*. In a preferred embodiment, bacterial host cells are used to express the

5 nucleic acid molecules of the instant invention. Useful expression vectors for bacterial hosts include bacterial plasmids, such as those from *E. coli, Bacillus* or *Streptomyces*, including pBluescript, pGEX-2T, pUC vectors, col E1, pCR1, pBR322, pMB9 and their derivatives, wider host range plasmids, such as RP4, phage DNAs, *e.g.*, the numerous derivatives of phage lambda, *e.g.*, NM989, λGT10 and λGT11, and other phages, *e.g.*,

10 M13 and filamentous single stranded phage DNA. Where *E. coli* is used as host, selectable markers are, analogously, chosen for selectivity in gram negative bacteria: *e.g.*, typical markers confer resistance to antibiotics, such as ampicillin, tetracycline, chloramphenicol, kanamycin, streptomycin and zeocin; auxotrophic markers can also be used.

15 In other embodiments, eukaryotic host cells, such as yeast, insect, mammalian or plant cells, may be used. Yeast cells, typically *S. cerevisiae*, are useful for eukaryotic genetic studies, due to the ease of targeting genetic changes by homologous recombination and the ability to easily complement genetic defects using recombinantly expressed proteins. Yeast cells are useful for identifying interacting protein components, *e.g.*

20 through use of a two-hybrid system. In a preferred embodiment, yeast cells are useful for protein expression. Vectors of the present invention for use in yeast will typically, but not invariably, contain an origin of replication suitable for use in yeast and a selectable marker that is functional in yeast. Yeast vectors include Yeast Integrating plasmids (*e.g.*, YIp5) and Yeast Replicating plasmids (the YRp and YEp series plasmids), Yeast Centromere

25 plasmids (the YCp series plasmids), Yeast Artificial Chromosomes (YACs) which are based on yeast linear plasmids, denoted YLp, pGPD-2, 2μ plasmids and derivatives thereof, and improved shuttle vectors such as those described in Gietz *et al.*, *Gene*, 74: 527-34 (1988) (YIplac, YEplac and YCplac). Selectable markers in yeast vectors include a variety of auxotrophic markers, the most common of which are (in *Saccharomyces*

30 *cerevisiae*) URA3, HIS3, LEU2, TRP1 and LYS2, which complement specific auxotrophic mutations, such as ura3-52, his3-D1, leu2-D1, trp1-D1 and lys2-201.

Insect cells may be chosen for high efficiency protein expression. Where the host cells are from *Spodoptera frugiperda, e.g.*, Sf9 and Sf21 cell lines, and expresSF™ cells

(Protein Sciences Corp., Meriden, CT, USA), the vector replicative strategy is typically based upon the baculovirus life cycle. Typically, baculovirus transfer vectors are used to replace the wild-type AcMNPV polyhedrin gene with a heterologous gene of interest. Sequences that flank the polyhedrin gene in the wild-type genome are positioned 5' and 3'

5   of the expression cassette on the transfer vectors. Following co-transfection with AcMNPV DNA, a homologous recombination event occurs between these sequences resulting in a recombinant virus carrying the gene of interest and the polyhedrin or p10 promoter. Selection can be based upon visual screening for lacZ fusion activity.

The host cells may also be mammalian cells, which are particularly useful for

10   expression of proteins intended as pharmaceutical agents, and for screening of potential agonists and antagonists of a protein or a physiological pathway. Mammalian vectors intended for autonomous extrachromosomal replication will typically include a viral origin, such as the SV40 origin (for replication in cell lines expressing the large T-antigen, such as COS1 and COS7 cells), the papillomavirus origin, or the EBV origin for long term

15   episomal replication (for use, *e.g.*, in 293-EBNA cells, which constitutively express the EBV EBNA-1 gene product and adenovirus E1A). Vectors intended for integration, and thus replication as part of the mammalian chromosome, can, but need not, include an origin of replication functional in mammalian cells, such as the SV40 origin. Vectors based upon viruses, such as adenovirus, adeno-associated virus, vaccinia virus, and

20   various mammalian retroviruses, will typically replicate according to the viral replicative strategy. Selectable markers for use in mammalian cells include, include but are not limited to, resistance to neomycin (G418), blasticidin, hygromycin and zeocin, and selection based upon the purine salvage pathway using HAT medium.

Expression in mammalian cells can be achieved using a variety of plasmids,

25   including pSV2, pBC12BI, and p91023, as well as lytic virus vectors (*e.g.*, vaccinia virus, adeno virus, and baculovirus), episomal virus vectors (*e.g.*, bovine papillomavirus), and retroviral vectors (*e.g.*, murine retroviruses). Useful vectors for insect cells include baculoviral vectors and pVL 941.

Plant cells can also be used for expression, with the vector replicon typically

30   derived from a plant virus (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) and selectable markers chosen for suitability in plants.

It is known that codon usage of different host cells may be different. For example, a plant cell and a human cell may exhibit a difference in codon preference for encoding a

particular amino acid. As a result, human mRNA may not be efficiently translated in a plant, bacteria or insect host cell. Therefore, another embodiment of this invention is directed to codon optimization. The codons of the nucleic acid molecules of the invention may be modified to resemble, as much as possible, genes naturally contained within the host cell without altering the amino acid sequence encoded by the nucleic acid molecule.

Any of a wide variety of expression control sequences may be used in these vectors to express the nucleic acid molecules of this invention. Such useful expression control sequences include the expression control sequences associated with structural genes of the foregoing expression vectors. Expression control sequences that control transcription include, e.g., promoters, enhancers and transcription termination sites. Expression control sequences in eukaryotic cells that control post-transcriptional events include splice donor and acceptor sites and sequences that modify the half-life of the transcribed RNA, e.g., sequences that direct poly(A) addition or binding sites for RNA-binding proteins. Expression control sequences that control translation include ribosome binding sites, sequences which direct targeted expression of the polypeptide to or within particular cellular compartments, and sequences in the 5' and 3' untranslated regions that modify the rate or efficiency of translation.

Examples of useful expression control sequences for a prokaryote, e.g., E. coli, will include a promoter, often a phage promoter, such as phage lambda pL promoter, the trc promoter, a hybrid derived from the trp and lac promoters, the bacteriophage T7 promoter (in E. coli cells engineered to express the T7 polymerase), the TAC or TRC system, the major operator and promoter regions of phage lambda, the control regions of fd coat protein, and the araBAD operon. Prokaryotic expression vectors may further include transcription terminators, such as the aspA terminator, and elements that facilitate translation, such as a consensus ribosome binding site and translation termination codon, Schomer et al., Proc. Natl. Acad. Sci. USA 83: 8506-8510 (1986).

Expression control sequences for yeast cells, typically S. cerevisiae, will include a yeast promoter, such as the CYC1 promoter, the GAL1 promoter, the GAL10 promoter, ADH1 promoter, the promoters of the yeast α-mating system, or the GPD promoter, and will typically have elements that facilitate transcription termination, such as the transcription termination signals from the CYC1 or ADH1 gene.

Expression vectors useful for expressing proteins in mammalian cells will include a promoter active in mammalian cells. These promoters include, but are not limited to,

those derived from mammalian viruses, such as the enhancer-promoter sequences from the immediate early gene of the human cytomegalovirus (CMV), the enhancer-promoter sequences from the Rous sarcoma virus long terminal repeat (RSV LTR), the enhancer-promoter from SV40 and the early and late promoters of adenovirus. Other expression

5   control sequences include the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase. Other expression control sequences include those-from the gene comprising the OSNA of interest. Often, expression is enhanced by incorporation of polyadenylation sites, such as the late SV40 polyadenylation site and the polyadenylation signal and transcription termination sequences from the bovine growth

10   hormone (BGH) gene, and ribosome binding sites. Furthermore, vectors can include introns, such as intron II of rabbit β-globin gene and the SV40 splice elements.

         Preferred nucleic acid vectors also include a selectable or amplifiable marker gene and means for amplifying the copy number of the gene of interest. Such marker genes are well known in the art. Nucleic acid vectors may also comprise stabilizing sequences (e.g.,

15   ori- or ARS-like sequences and telomere-like sequences), or may alternatively be designed to favor directed or non-directed integration into the host cell genome. In a preferred embodiment, nucleic acid sequences of this invention are inserted in frame into an expression vector that allows a high level expression of an RNA which encodes a protein comprising the encoded nucleic acid sequence of interest. Nucleic acid cloning and

20   sequencing methods are well known to those of skill in the art and are described in an assortment of laboratory manuals, including Sambrook (1989), *supra*, Sambrook (2000), *supra*; and Ausubel (1992), *supra*, Ausubel (1999), *supra*. Product information from manufacturers of biological, chemical and immunological reagents also provide useful information.

25   Expression vectors may be either constitutive or inducible. Inducible vectors include either naturally inducible promoters, such as the trc promoter, which is regulated by the lac operon, and the pL promoter, which is regulated by tryptophan, the MMTV-LTR promoter, which is inducible by dexamethasone, or can contain synthetic promoters and/or additional elements that confer inducible control on adjacent promoters.

30   Examples of inducible synthetic promoters are the hybrid Plac/ara-1 promoter and the PLtetO-1 promoter. The PLtetO-1 promoter takes advantage of the high expression levels from the PL promoter of phage lambda, but replaces the lambda repressor sites with two copies of operator 2 of the Tn10 tetracycline resistance operon, causing this promoter to

be tightly repressed by the Tet repressor protein and induced in response to tetracycline (Tc) and Tc derivatives such as anhydrotetracycline. Vectors may also be inducible because they contain hormone response elements, such as the glucocorticoid response element (GRE) and the estrogen response element (ERE), which can confer hormone

5   inducibility where vectors are used for expression in cells having the respective hormone receptors. To reduce background levels of expression, elements responsive to ecdysone, an insect hormone, can be used instead, with coexpression of the ecdysone receptor.

In one embodiment of the invention, expression vectors can be designed to fuse the expressed polypeptide to small protein tags that facilitate purification and/or visualization.

10   Such tags include a polyhistidine tag that facilitates purification of the fusion protein by immobilized metal affinity chromatography, for example using NiNTA resin (Qiagen Inc., Valencia, CA, USA) or TALON™ resin (cobalt immobilized affinity chromatography medium, Clontech Labs, Palo Alto, CA, USA). The fusion protein can include a chitin-binding tag and self-excising intein, permitting chitin-based purification with self-removal

15   of the fused tag (IMPACT™ system, New England Biolabs, Inc., Beverley, MA, USA). Alternatively, the fusion protein can include a calmodulin-binding peptide tag, permitting purification by calmodulin affinity resin (Stratagene, La Jolla, CA, USA), or a specifically excisable fragment of the biotin carboxylase carrier protein, permitting purification of *in vivo* biotinylated protein using an avidin resin and subsequent tag removal (Promega,

20   Madison, WI, USA). As another useful alternative, the polypeptides of the present invention can be expressed as a fusion to glutathione-S-transferase, the affinity and specificity of binding to glutathione permitting purification using glutathione affinity resins, such as Glutathione-Superflow Resin (Clontech Laboratories, Palo Alto, CA, USA), with subsequent elution with free glutathione. Other tags include, for example, the

25   Xpress epitope, detectable by anti-Xpress antibody (Invitrogen, Carlsbad, CA, USA), a myc tag, detectable by anti-myc tag antibody, the V5 epitope, detectable by anti-V5 antibody (Invitrogen, Carlsbad, CA, USA), FLAG® epitope, detectable by anti-FLAG® antibody (Stratagene, La Jolla, CA, USA), and the HA epitope, detectable by anti-HA antibody.

30   For secretion of expressed polypeptides, vectors can include appropriate sequences that encode secretion signals, such as leader peptides. For example, the pSecTag2 vectors (Invitrogen, Carlsbad, CA, USA) are 5.2 kb mammalian expression vectors that carry the

secretion signal from the V-J2-C region of the mouse Ig kappa-chain for efficient secretion of recombinant proteins from a variety of mammalian cell lines.

Expression vectors can also be designed to fuse proteins encoded by the heterologous nucleic acid insert to polypeptides that are larger than purification and/or
5   identification tags. Useful protein fusions include those that permit display of the encoded protein on the surface of a phage or cell, fusions to intrinsically fluorescent proteins, such as those that have a green fluorescent protein (GFP)-like chromophore, fusions to the IgG Fc region, and fusions for use in two hybrid systems.

Vectors for phage display fuse the encoded polypeptide to, *e.g.*, the gene III
10   protein (pIII) or gene VIII protein (pVIII) for display on the surface of filamentous phage, such as M13. *See* Barbas *et al.*, Phage Display: A Laboratory Manual, Cold Spring Harbor Laboratory Press (2001); Kay *et al.* (eds.), Phage Display of Peptides and Proteins: A Laboratory Manual, Academic Press, Inc., (1996); Abelson *et al.* (eds.), Combinatorial Chemistry (Methods in Enzymology, Vol. 267) Academic Press (1996). Vectors for yeast
15   display, *e.g.* the pYD1 yeast display vector (Invitrogen, Carlsbad, CA, USA), use the α-agglutinin yeast adhesion receptor to display recombinant protein on the surface of *S. cerevisiae*. Vectors for mammalian display, *e.g.*, the pDisplay™ vector (Invitrogen, Carlsbad, CA, USA), target recombinant proteins using an N-terminal cell surface targeting signal and a C-terminal transmembrane anchoring domain of platelet derived
20   growth factor receptor.

A wide variety of vectors now exist that fuse proteins encoded by heterologous nucleic acids to the chromophore of the substrate-independent, intrinsically fluorescent green fluorescent protein from *Aequorea victoria* ("GFP") and its variants. The GFP-like chromophore can be selected from GFP-like chromophores found in naturally occurring
25   proteins, such as *A. victoria* GFP (GenBank accession number AAA27721), *Renilla reniformis* GFP, FP583 (GenBank accession no. AF168419) (DsRed), FP593 (AF272711), FP483 (AF168420), FP484 (AF168424), FP595 (AF246709), FP486 (AF168421), FP538 (AF168423), and FP506 (AF168422), and need include only so much of the native protein as is needed to retain the chromophore's intrinsic fluorescence. Methods for determining
30   the minimal domain required for fluorescence are known in the art. *See* Li *et al.*, *J. Biol. Chem.* 272: 28545-28549 (1997). Alternatively, the GFP-like chromophore can be selected from GFP-like chromophores modified from those found in nature. The methods for engineering such modified GFP-like chromophores and testing them for fluorescence

activity, both alone and as part of protein fusions, are well known in the art. *See* Heim *et al.*, *Curr. Biol.* 6: 178-182 (1996) and Palm *et al.*, *Methods Enzymol.* 302: 378-394 (1999). A variety of such modified chromophores are now commercially available and can readily be used in the fusion proteins of the present invention. These include EGFP ("enhanced

5   GFP"), EBFP ("enhanced blue fluorescent protein"), BFP2, EYFP ("enhanced yellow fluorescent protein"), ECFP ("enhanced cyan fluorescent protein") or Citrine. EGFP (*see, e.g,* Cormack *et al.*, *Gene* 173: 33–38 (1996); U.S. Patent Nos. 6,090,919 and 5,804,387, the disclosures of which are incorporated herein by reference in their entireties) is found on a variety of vectors, both plasmid and viral, which are available commercially

10   (Clontech Labs, Palo Alto, CA, USA); EBFP is optimized for expression in mammalian cells whereas BFP2, which retains the original jellyfish codons, can be expressed in bacteria (*see, e.g,.* Heim *et al.*, *Curr. Biol.* 6: 178-182 (1996) and Cormack *et al.*, *Gene* 173: 33-38 (1996)). Vectors containing these blue-shifted variants are available from Clontech Labs (Palo Alto, CA, USA). Vectors containing EYFP, ECFP (*see, e.g.*, Heim *et*

15   *al.*, *Curr. Biol.* 6: 178-182 (1996); Miyawaki *et al.*, *Nature* 388: 882-887 (1997)) and Citrine (*see, e.g.*, Heikal *et al.*, *Proc. Natl. Acad. Sci. USA* 97: 11996-12001 (2000)) are also available from Clontech Labs. The GFP-like chromophore can also be drawn from other modified GFPs, including those described in U.S. Patent Nos. 6,124,128; 6,096,865; 6,090,919; 6,066,476; 6,054,321; 6,027,881; 5,968,750; 5,874,304; 5,804,387; 5,777,079;

20   5,741,668; and 5,625,048, the disclosures of which are incorporated herein by reference in their entireties. *See* also Conn (ed.), Green Fluorescent Protein (Methods in Enzymology, Vol. 302), Academic Press, Inc. (1999); Yang, *et al.*, *J Biol Chem*, 273: 8212-6 (1998); Bevis *et al.*, *Nature Biotechnology*, 20:83-7 (2002). The GFP-like chromophore of each of these GFP variants can usefully be included in the fusion proteins of the present

25   invention.

Fusions to the IgG Fc region increase serum half-life of protein pharmaceutical products through interaction with the FcRn receptor (also denominated the FcRp receptor and the Brambell receptor, FcRb), further described in International Patent Application nos. WO 97/43316, WO 97/34631, WO 96/32478, WO 96/18412, the disclosures of which

30   are incorporated herein by reference in their entireties.

For long-term, high-yield recombinant production of the polypeptides of the present invention, stable expression is preferred. Stable expression is readily achieved by integration into the host cell genome of vectors having selectable markers, followed by

selection of these integrants. Vectors such as pUB6/V5-His A, B, and C (Invitrogen, Carlsbad, CA, USA) are designed for high-level stable expression of heterologous proteins in a wide range of mammalian tissue types and cell lines. pUB6/V5-His uses the promoter/enhancer sequence from the human ubiquitin C gene to drive expression of

5   recombinant proteins: expression levels in 293, CHO, and NIH3T3 cells are comparable to levels from the CMV and human EF-1a promoters. The bsd gene permits rapid selection of stably transfected mammalian cells with the potent antibiotic blasticidin.

Replication incompetent retroviral vectors, typically derived from Moloney murine leukemia virus, also are useful for creating stable transfectants having integrated provirus.

10   The highly efficient transduction machinery of retroviruses, coupled with the availability of a variety of packaging cell lines such as RetroPack™ PT 67, EcoPack2™-293, AmphoPack-293, and GP2-293 cell lines (all available from Clontech Laboratories, Palo Alto, CA, USA) allow a wide host range to be infected with high efficiency; varying the multiplicity of infection readily adjusts the copy number of the integrated provirus.

15   Of course, not all vectors and expression control sequences will function equally well to express the nucleic acid molecules of this invention. Neither will all hosts function equally well with the same expression system. However, one of skill in the art may make a selection among these vectors, expression control sequences and hosts without undue experimentation and without departing from the scope of this invention. For example, in

20   selecting a vector, the host must be considered because the vector must be replicated in it. The vector's copy number, the ability to control that copy number, the ability to control integration, if any, and the expression of any other proteins encoded by the vector, such as antibiotic or other selection markers, should also be considered. The present invention further includes host cells comprising the vectors of the present invention, either present

25   episomally within the cell or integrated, in whole or in part, into the host cell chromosome. Among other considerations, some of which are described above, a host cell strain may be chosen for its ability to process the expressed polypeptide in the desired fashion. Such post-translational modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation, and it

30   is an aspect of the present invention to provide OSPs with such post-translational modifications.

In selecting an expression control sequence, a variety of factors should also be considered. These include, for example, the relative strength of the sequence, its

controllability, and its compatibility with the nucleic acid molecules of this invention, particularly with regard to potential secondary structures. Unicellular hosts should be selected by consideration of their compatibility with the chosen vector, the toxicity of the product coded for by the nucleic acid sequences of this invention, their secretion

5   characteristics, their ability to fold the polypeptide correctly, their fermentation or culture requirements, and the ease of purification from them of the products coded for by the nucleic acid molecules of this invention.

The recombinant nucleic acid molecules and more particularly, the expression vectors of this invention may be used to express the polypeptides of this invention as

10   recombinant polypeptides in a heterologous host cell. The polypeptides of this invention may be full-length or less than full-length polypeptide fragments recombinantly expressed from the nucleic acid molecules according to this invention. Such polypeptides include analogs, derivatives and muteins that may or may not have biological activity.

Vectors of the present invention will also often include elements that permit *in*

15   *vitro* transcription of RNA from the inserted heterologous nucleic acid. Such vectors typically include a phage promoter, such as that from T7, T3, or SP6, flanking the nucleic acid insert. Often two different such promoters flank the inserted nucleic acid, permitting separate *in vitro* production of both sense and antisense strands.

Transformation and other methods of introducing nucleic acids into a host cell

20   (*e.g.*, conjugation, protoplast transformation or fusion, transfection, electroporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion) can be accomplished by a variety of methods which are well known in the art (*See*, for instance, Ausubel, *supra*, and Sambrook *et al.*, *supra*). Bacterial, yeast, plant or mammalian cells are transformed or transfected with an

25   expression vector, such as a plasmid, a cosmid, or the like, wherein the expression vector comprises the nucleic acid of interest. Alternatively, the cells may be infected by a viral expression vector comprising the nucleic acid of interest. Depending upon the host cell, vector, and method of transformation used, transient or stable expression of the polypeptide will be constitutive or inducible. One having ordinary skill in the art will be

30   able to decide whether to express a polypeptide transiently or stably, and whether to express the protein constitutively or inducibly.

A wide variety of unicellular host cells are useful in expressing the DNA sequences of this invention. These hosts may include well known eukaryotic and

prokaryotic hosts, such as strains of, fungi, yeast, insect cells such as *Spodoptera frugiperda* (SF9), animal cells such as CHO, as well as plant cells in tissue culture. Representative examples of appropriate host cells include, but are not limited to, bacterial cells, such as *E. coli*, *Caulobacter crescentus*, *Streptomyces* species, and *Salmonella*

5   *typhimurium*; yeast cells, such as *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris*, *Pichia methanolica*; insect cell lines, such as those from *Spodoptera frugiperda* — *e.g.*, Sf9 and Sf21 cell lines, and expresSF™ cells (Protein Sciences Corp., Meriden, CT, USA) — *Drosophila* S2 cells, and *Trichoplusia ni* High Five® Cells (Invitrogen, Carlsbad, CA, USA); and mammalian cells. Typical mammalian cells include

10   BHK cells, BSC 1 cells, BSC 40 cells, BMT 10 cells, VERO cells, COS1 cells, COS7 cells, Chinese hamster ovary (CHO) cells, 3T3 cells, NIH 3T3 cells, 293 cells, HEPG2 cells, HeLa cells, L cells, MDCK cells, HEK293 cells, WI38 cells, murine ES cell lines (*e.g.*, from strains 129/SV, C57/BL6, DBA-1, 129/SVJ), K562 cells, Jurkat cells, and BW5147 cells. Other mammalian cell lines are well known and readily available from

15   the American Type Culture Collection (ATCC) (Manassas, VA, USA) and the National Institute of General Medical Sciences (NIGMS) Human Genetic Cell Repository at the Coriell Cell Repositories (Camden, NJ, USA). Cells or cell lines derived from ovarian are particularly preferred because they may provide a more native post-translational processing. Particularly preferred are human ovarian cells.

20   Particular details of the transfection, expression and purification of recombinant proteins are well documented and are understood by those of skill in the art. Further details on the various technical aspects of each of the steps used in recombinant production of foreign genes in bacterial cell expression systems can be found in a number of texts and laboratory manuals in the art. *See, e.g.*, Ausubel (1992), *supra*, Ausubel

25   (1999), *supra*, Sambrook (1989), *supra*, and Sambrook (2001), *supra*.

       Methods for introducing the vectors and nucleic acid molecules of the present invention into the host cells are well known in the art; the choice of technique will depend primarily upon the specific vector to be introduced and the host cell chosen.

       Nucleic acid molecules and vectors may be introduced into prokaryotes, such as *E.*

30   *coli*, in a number of ways. For instance, phage lambda vectors will typically be packaged using a packaging extract (*e.g.*, Gigapack® packaging extract, Stratagene, La Jolla, CA, USA), and the packaged virus used to infect *E. coli*.

Plasmid vectors will typically be introduced into chemically competent or electrocompetent bacterial cells. *E. coli* cells can be rendered chemically competent by treatment, *e.g.*, with CaCl₂, or a solution of Mg^{2+}, Mn^{2+}, Ca^{2+}, Rb^+ or K^+, dimethyl sulfoxide, dithiothreitol, and hexamine cobalt (III), Hanahan, *J. Mol. Biol.* 166(4):557-80

5 (1983), and vectors introduced by heat shock. A wide variety of chemically competent strains are also available commercially (*e.g.*, Epicurian Coli® XL10-Gold® Ultracompetent Cells (Stratagene, La Jolla, CA, USA); DH5α competent cells (Clontech Laboratories, Palo Alto, CA, USA); and TOP10 Chemically Competent E. coli Kit (Invitrogen, Carlsbad, CA, USA)). Bacterial cells can be rendered electrocompetent to

10 take up exogenous DNA by electroporation by various pre-pulse treatments; vectors are introduced by electroporation followed by subsequent outgrowth in selected media. An extensive series of protocols is provided by BioRad (Richmond, CA, USA).

Vectors can be introduced into yeast cells by spheroplasting, treatment with lithium salts, electroporation, or protoplast fusion. Spheroplasts are prepared by the action

15 of hydrolytic enzymes such as a snail-gut extract, usually denoted Glusulase or Zymolyase, or an enzyme from *Arthrobacter luteus* to remove portions of the cell wall in the presence of osmotic stabilizers, typically 1 M sorbitol. DNA is added to the spheroplasts, and the mixture is co-precipitated with a solution of polyethylene glycol (PEG) and Ca^{2+}. Subsequently, the cells are resuspended in a solution of sorbitol, mixed

20 with molten agar and then layered on the surface of a selective plate containing sorbitol.

For lithium-mediated transformation, yeast cells are treated with lithium acetate to permeabilize the cell wall, DNA is added and the cells are co-precipitated with PEG. The cells are exposed to a brief heat shock, washed free of PEG and lithium acetate, and subsequently spread on plates containing ordinary selective medium. Increased

25 frequencies of transformation are obtained by using specially-prepared single-stranded carrier DNA and certain organic solvents. Schiestl *et al.*, *Curr. Genet.* 16(5-6): 339-46 (1989).

For electroporation, freshly-grown yeast cultures are typically washed, suspended in an osmotic protectant, such as sorbitol, mixed with DNA, and the cell suspension

30 pulsed in an electroporation device. Subsequently, the cells are spread on the surface of plates containing selective media. Becker *et al.*, *Methods Enzymol.* 194: 182-187 (1991). The efficiency of transformation by electroporation can be increased over 100-fold by

using PEG, single-stranded carrier DNA and cells that are in late log-phase of growth.

Larger constructs, such as YACs, can be introduced by protoplast fusion.

Mammalian and insect cells can be directly infected by packaged viral vectors, or

transfected by chemical or electrical means. For chemical transfection, DNA can be

5   coprecipitated with $CaPO_4$ or introduced using liposomal and nonliposomal lipid-based

agents. Commercial kits are available for $CaPO_4$ transfection (CalPhos™ Mammalian

Transfection Kit, Clontech Laboratories, Palo Alto, CA, USA), and lipid-mediated

transfection can be practiced using commercial reagents, such as LIPOFECTAMINE™

2000, LIPOFECTAMINE™ Reagent, CELLFECTIN® Reagent, and LIPOFECTIN®

10   Reagent (Invitrogen, Carlsbad, CA, USA), DOTAP Liposomal Transfection Reagent,

FuGENE 6, X-tremeGENE Q2, DOSPER, (Roche Molecular Biochemicals, Indianapolis,

IN USA), Effectene™, PolyFect®, Superfect® (Qiagen, Inc., Valencia, CA, USA).

Protocols for electroporating mammalian cells can be found in, for example, ; Norton *et*

al. (eds.), Gene Transfer Methods: Introducing DNA into Living Cells and Organisms,

15   BioTechniques Books, Eaton Publishing Co. (2000). Other transfection techniques

include transfection by particle bombardment and microinjection. *See, e.g.*, Cheng *et al.*,

Proc. Natl. Acad. Sci. USA 90(10): 4455-9 (1993); Yang *et al.*, *Proc. Natl. Acad. Sci. USA*

87(24): 9568-72 (1990).

Production of the recombinantly produced proteins of the present invention can

20   optionally be followed by purification.

Purification of recombinantly expressed proteins is now well within the skill in the

art and thus need not be detailed here. *See, e.g.*, Thorner *et al.* (eds.), Applications of

Chimeric Genes and Hybrid Proteins, Part A: Gene Expression and Protein Purification

(Methods in Enzymology, Vol. 326), Academic Press (2000); Harbin (ed.), Cloning, Gene

25   Expression and Protein Purification : Experimental Procedures and Process Rationale,

Oxford Univ. Press (2001); Marshak *et al.*, Strategies for Protein Purification and

Characterization: A Laboratory Course Manual, Cold Spring Harbor Laboratory Press

(1996); and Roe (ed.), Protein Purification Applications, Oxford University Press (2001).

Briefly, however, if purification tags have been fused through use of an expression

30   vector that appends such tag, purification can be effected, at least in part, by means

appropriate to the tag, such as use of immobilized metal affinity chromatography for

polyhistidine tags. Other techniques common in the art include ammonium sulfate

fractionation, immunoprecipitation, fast protein liquid chromatography (FPLC), high performance liquid chromatography (HPLC), and preparative gel electrophoresis.

<u>Polypeptides, including Fragments Muteins, Homologous Proteins, Allelic Variants, Analogs and Derivatives</u>

5   Another aspect of the invention relates to polypeptides encoded by the nucleic acid molecules described herein. In a preferred embodiment, the polypeptide is a ovarian specific polypeptide (OSP). In an even more preferred embodiment, the polypeptide comprises an amino acid sequence of SEQ ID NO:100-192 or is derived from a polypeptide having the amino acid sequence of SEQ ID NO: 100-192. A polypeptide as

10   defined herein may be produced recombinantly, as discussed *supra*, may be isolated from . a cell that naturally expresses the protein, or may be chemically synthesized following the teachings of the specification and using methods well known to those having ordinary skill in the art.

         Polypeptides of the present invention may also comprise a part or fragment of a

15   OSP. In a preferred embodiment, the fragment is derived from a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 100-192. Polypeptides of the present invention comprising a part or fragment of an entire OSP may or may not be OSPs. For example, a full-length polypeptide may be ovarian-specific, while a fragment thereof may be found in other tissues as well as in ovarian. A

20   polypeptide that is not a OSP, whether it is a fragment, analog, mutein, homologous protein or derivative, is nevertheless useful, especially for immunizing animals to prepare anti-OSP antibodies. In a preferred embodiment, the part or fragment is a OSP. Methods of determining whether a polypeptide of the present invention is a OSP are described *infra*.

25   Polypeptides of the present invention comprising fragments of at least 6 contiguous amino acids are also useful in mapping B cell and T cell epitopes of the reference protein. *See, e.g.*, Geysen *et al.*, *Proc. Natl. Acad. Sci. USA* 81: 3998-4002 (1984) and U.S. Patent Nos. 4,708,871 and 5,595,915, the disclosures of which are incorporated herein by reference in their entireties. Because the fragment need not itself

30   be immunogenic, part of an immunodominant epitope, nor even recognized by native antibody, to be useful in such epitope mapping, all fragments of at least 6 amino acids of a polypeptide of the present invention have utility in such a study.

Polypeptides of the present invention comprising fragments of at least 8 contiguous amino acids, often at least 15 contiguous amino acids, are useful as immunogens for raising antibodies that recognize polypeptides of the present invention. *See, e.g.*, Lerner, Nature 299: 592-596 (1982); Shinnick *et al.*, *Annu. Rev. Microbiol.* 37:

5   425-46 (1983); Sutcliffe *et al.*, *Science* 219: 660-6 (1983). As further described in the above-cited references, virtually all 8-mers, conjugated to a carrier, such as a protein, prove immunogenic and are capable of eliciting antibody for the conjugated peptide; accordingly, all fragments of at least 8 amino acids of the polypeptides of the present invention have utility as immunogens.

10   Polypeptides comprising fragments of at least 8, 9, 10 or 12 contiguous amino acids are also useful as competitive inhibitors of binding of the entire polypeptide, or a portion thereof, to antibodies (as in epitope mapping), and to natural binding partners, such as subunits in a multimeric complex or to receptors or ligands of the subject protein; this competitive inhibition permits identification and separation of molecules that bind

15   specifically to the polypeptide of interest. See U.S. Patent Nos. 5,539,084 and 5,783,674, incorporated herein by reference in their entireties.

The polypeptide of the present invention thus preferably is at least 6 amino acids in length, typically at least 8, 9, 10 or 12 amino acids in length, and often at least 15 amino acids in length. Often, the polypeptide of the present invention is at least 20 amino acids

20   in length, even 25 amino acids, 30 amino acids, 35 amino acids, or 50 amino acids or more in length. Of course, larger polypeptides having at least 75 amino acids, 100 amino acids, or even 150 amino acids are also useful, and at times preferred.

One having ordinary skill in the art can produce fragments by truncating the nucleic acid molecule, *e.g.*, a OSNA, encoding the polypeptide and then expressing it

25   recombinantly. Alternatively, one can produce a fragment by chemically synthesizing a portion of the full-length polypeptide. One may also produce a fragment by enzymatically cleaving either a recombinant polypeptide or an isolated naturally occurring polypeptide. Methods of producing polypeptide fragments are well known in the art. *See, e.g.*, Sambrook (1989), *supra;* Sambrook (2001), *supra;* Ausubel (1992), *supra;* and Ausubel

30   (1999), *supra.* In one embodiment, a polypeptide comprising only a fragment, preferably a fragment of a OSP, may be produced by chemical or enzymatic cleavage of a OSP polypeptide. In a preferred embodiment, a polypeptide fragment is produced by

expressing a nucleic acid molecule of the present invention encoding a fragment, preferably of a OSP, in a host cell.

Polypeptides of the present invention are also inclusive of mutants, fusion proteins, homologous proteins and allelic variants.

5   A mutant protein, or mutein, may have the same or different properties compared to a naturally occurring polypeptide and comprises at least one amino acid insertion, duplication, deletion, rearrangement or substitution compared to the amino acid sequence of a native polypeptide. Small deletions and insertions can often be found that do not alter the function of a protein. Muteins may or may not be ovarian-specific. Preferably, the

10   mutein is ovarian-specific. More preferably the mutein is a polypeptide that comprises at least one amino acid insertion, duplication, deletion, rearrangement or substitution compared to the amino acid sequence of SEQ ID NO: 100-192. Accordingly, in a preferred embodiment, the mutein is one that exhibits at least 50% sequence identity, more preferably at least 60% sequence identity, even more preferably at least 70%, yet more

15   preferably at least 80% sequence identity to a OSP comprising an amino acid sequence of SEQ ID NO: 100-192. In a yet more preferred embodiment, the mutein exhibits at least 85%, more preferably 90%, even more preferably 95% or 96%, and yet more preferably at least 97%, 98%, 99% or 99.5% sequence identity to a OSP comprising an amino acid sequence of SEQ ID NO: 100-192.

20   A mutein may be produced by isolation from a naturally occurring mutant cell, tissue or organism. A mutein may be produced by isolation from a cell, tissue or organism that has been experimentally mutagenized. Alternatively, a mutein may be produced by chemical manipulation of a polypeptide, such as by altering the amino acid residue to another amino acid residue using synthetic or semi-synthetic chemical techniques. In a

25   preferred embodiment, a mutein is produced from a host cell comprising a mutated nucleic acid molecule compared to the naturally occurring nucleic acid molecule. For instance, one may produce a mutein of a polypeptide by introducing one or more mutations into a nucleic acid molecule of the invention and then expressing it recombinantly. These mutations may be targeted, in which particular encoded amino acids are altered, or may be

30   untargeted, in which random encoded amino acids within the polypeptide are altered. Muteins with random amino acid alterations can be screened for a particular biological activity or property, particularly whether the polypeptide is ovarian-specific, as described below. Multiple random mutations can be introduced into the gene by methods well

known to the art, *e.g.*, by error-prone PCR, shuffling, oligonucleotide-directed

mutagenesis, assembly PCR, sexual PCR mutagenesis, *in vivo* mutagenesis, cassette

mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis and site-

specific mutagenesis. Methods of producing muteins with targeted or random amino acid

5   alterations are well known in the art. *See, e.g.*, Sambrook (1989), *supra;* Sambrook

(2001), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), as well as U.S. Patent No.

5,223,408, which is herein incorporated by reference in its entirety.

The invention also contemplates polypeptides that are homologous to a

polypeptide of the invention. In a preferred embodiment, the polypeptide is homologous

10   to a OSP. In an even more preferred embodiment, the polypeptide is homologous to a

OSP selected from the group having an amino acid sequence of SEQ ID NO: 100-192. By

homologous polypeptide it is means one that exhibits significant sequence identity to a

OSP, preferably a OSP having an amino acid sequence of SEQ ID NO: 100-192. By

significant sequence identity it is meant that the homologous polypeptide exhibits at least

15   50% sequence identity, more preferably at least 60% sequence identity, even more

preferably at least 70%, yet more preferably at least 80% sequence identity to a OSP

comprising an amino acid sequence of SEQ ID NO: 100-192. More preferred are

homologous polypeptides exhibiting at least 85%, more preferably 90%, even more

preferably 95% or 96%, and yet more preferably at least 97% or 98% sequence identity to

20   a OSP comprising an amino acid sequence of SEQ ID NO: 100-192. Most preferably, the

homologous polypeptide exhibits at least 99%, more preferably 99.5%, even more

preferably 99.6%, 99.7%, 99.8% or 99.9% sequence identity to a OSP comprising an

amino acid sequence of SEQ ID NO: 100-192. In a preferred embodiment, the amino acid

substitutions of the homologous polypeptide are conservative amino acid substitutions as

25   discussed above.

Homologous polypeptides of the present invention also comprise polypeptide

encoded by a nucleic acid molecule that selectively hybridizes to a OSNA or an antisense

sequence thereof. In this embodiment, it is preferred that the homologous polypeptide be

encoded by a nucleic acid molecule that hybridizes to a OSNA under low stringency,

30   moderate stringency or high stringency conditions, as defined herein. More preferred is a

homologous polypeptide encoded by a nucleic acid sequence which hybridizes to a OSNA

selected from the group consisting of SEQ ID NO: 1-99 or a homologous polypeptide

encoded by a nucleic acid molecule that hybridizes to a nucleic acid molecule that encodes

a OSP, preferably an OSP of SEQ ID NO:100-192 under low stringency, moderate

stringency or high stringency conditions, as defined herein.

Homologous polypeptides of the present invention may be naturally occurring and

derived from another species, especially one derived from another primate, such as

5   chimpanzee, gorilla, rhesus macaque, or baboon, wherein the homologous polypeptide

comprises an amino acid sequence that exhibits significant sequence identity to that of

SEQ ID NO: 100-192. The homologous polypeptide may also be a naturally occurring

polypeptide from a human, when the OSP is a member of a family of polypeptides. The

homologous polypeptide may also be a naturally occurring polypeptide derived from a

10   non-primate, mammalian species, including without limitation, domesticated species, e.g.,

dog, cat, mouse, rat, rabbit, guinea pig, hamster, cow, horse, goat or pig. The homologous

polypeptide may also be a naturally occurring polypeptide derived from a non-mammalian

species, such as birds or reptiles. The naturally occurring homologous protein may be

isolated directly from humans or other species. Alternatively, the nucleic acid molecule

15   encoding the naturally occurring homologous polypeptide may be isolated and used to

express the homologous polypeptide recombinantly. The homologous polypeptide may

also be one that is experimentally produced by random mutation of a nucleic acid

molecule and subsequent expression of the nucleic acid molecule. Alternatively, the

homologous polypeptide may be one that is experimentally produced by directed mutation

20   of one or more codons to alter the encoded amino acid of a OSP. In a preferred

embodiment, the homologous polypeptide encodes a polypeptide that is a OSP.

Relatedness of proteins can also be characterized using a second functional test,

the ability of a first protein competitively to inhibit the binding of a second protein to an

antibody. It is, therefore, another aspect of the present invention to provide isolated

25   polpeptide not only identical in sequence to those described with particularity herein, but

also to provide isolated polypeptide ("cross-reactive proteins") that competitively inhibit

the binding of antibodies to all or to a portion of various of the isolated polypeptides of the

present invention. Such competitive inhibition can readily be determined using

immunoassays well known in the art.

30   As discussed above, single nucleotide polymorphisms (SNPs) occur frequently in

eukaryotic genomes, and the sequence determined from one individual of a species may

differ from other allelic forms present within the population. Thus, polypeptides of the

present invention are also inclusive of those encoded by an allelic variant of a nucleic acid

molecule encoding a OSP. In this embodiment, it is preferred that the polypeptide be encoded by an allelic variant of a gene that encodes a polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO: 100-192. More preferred is that the polypeptide be encoded by an allelic variant of a gene that has the nucleic acid

5   sequence selected from the group consisting of SEQ ID NO: 1-99.

Polypeptides of the present invention are also inclusive of derivative polypeptides encoded by a nucleic acid molecule according to the instant invention. In this embodiment, it is preferred that the polypeptide be a OSP. Also preferred are derivative polypeptides having an amino acid sequence selected from the group consisting of SEQ

10   ID NO: 100-192 and which has been acetylated, carboxylated, phosphorylated, glycosylated, ubiquitinated or other PTMs. In another preferred embodiment, the derivative has been labeled with, e.g., radioactive isotopes such as ^{125}I, ^{32}P, ^{35}S, and ^{3}H. In another preferred embodiment, the derivative has been labeled with fluorophores, chemiluminescent agents, enzymes, and antiligands that can serve as specific binding pair

15   members for a labeled ligand.

Polypeptide modifications are well known to those of skill and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in

20   most basic texts, such as, for instance Creighton, Protein Structure and Molecular Properties, 2nd ed., W. H. Freeman and Company (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, in Johnson (ed.), Posttranslational Covalent Modification of Proteins, pgs. 1-12, Academic Press (1983); . Seifter et al., Meth. Enzymol. 182: 626-646 (1990) and Rattan et al., Ann. N.Y. Acad. Sci.

25   663: 48-62 (1992).

One may determine whether a polypeptide of the invention is likely to be post-translationally modified by analyzing the sequence of the polypeptide to determine if there are peptide motifs indicative of sites for post-translational modification. There are a number of computer programs that permit prediction of post-translational modifications.

30   See, e.g., www.expasy.org (accessed November 11, 2002), which includes PSORT, for prediction of protein sorting signals and localization sites, SignalP, for prediction of signal peptide cleavage sites, MITOPROT and Predotar, for prediction of mitochondrial targeting sequences, NetOGlyc, for prediction of type O-glycosylation sites in mammalian proteins,

big-PI Predictor and DGPI, for prediction of prenylation-anchor and cleavage sites, and NetPhos, for prediction of Ser, Thr and Tyr phosphorylation sites in eukaryotic proteins. Other computer programs, such as those included in GCG, also may be used to determine post-translational modification peptide motifs.

5   General examples of types of post-translational modifications include, but are not limited to: (Z)-dehydrobutyrine; 1-chondroitin sulfate-L-aspartic acid ester; 1'-glycosyl-L-tryptophan; 1'-phospho-L-histidine; 1-thioglycine; 2'-(S-L-cysteinyl)-L-histidine; 2'-[3-carboxamido (trimethylammonio)propyl]-L-histidine; 2'-alpha-mannosyl-L-tryptophan; 2-methyl-L-glutamine; 2-oxobutanoic acid; 2-pyrrolidone carboxylic acid; 3'-(1'-L-histidyl)-

10   L-tyrosine; 3'-(8alpha-FAD)-L-histidine; 3'-(S-L-cysteinyl)-L-tyrosine; 3', 3",5'-triiodo-L-thyronine; 3'-4'-phospho-L-tyrosine; 3-hydroxy-L-proline; 3'-methyl-L-histidine; 3-methyl-L-lanthionine; 3'-phospho-L-histidine; 4'-(L-tryptophan)-L-tryptophyl quinone; 42 N-cysteinyl-glycosylphosphatidylinositolethanolamine; 43 -(T-L-histidyl)-L-tyrosine; 4-hydroxy-L-arginine; 4-hydroxy-L-lysine; 4-hydroxy-L-proline; 5'-(N6-L-lysine)-L-

15   topaquinone; 5-hydroxy-L-lysine; 5-methyl-L-arginine; alpha-1-microglobulin-Ig alpha complex chromophore; bis-L-cysteinyl bis-L-histidino diiron disulfide; bis-L--cysteinyl-L-N3'-histidino-L-serinyl tetrairon' tetrasulfide; chondroitin sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-serine; D-alanine; D-allo-isoleucine; D-asparagine; dehydroalanine; dehydrotyrosine; dermatan 4-sulfate D-glucuronyl-D-galactosyl-D-

20   galactosyl-D-xylosyl-L-serine; D-glucuronyl-N-glycine; dipyrrolylmethanemethyl-L-cysteine; D-leucine; D-methionine; D-phenylalanine; D-serine; D-tryptophan; glycine amide; glycine oxazolecarboxylic acid; glycine thiazolecarboxylic acid; heme P450-bis-L-cysteine-L-tyrosine; heme-bis-L-cysteine; hemediol-L-aspartyl ester-L-glutamyl ester; hemediol-L-aspartyl ester-L-glutamyl ester-L-methionine sulfonium; heme-L-cysteine;

25   heme-L-histidine; heparan sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-serine; heme P450-bis-L-cysteine-L-lysine; hexakis-L-cysteinyl hexairon hexasulfide; keratan sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-threonine; L oxoalanine- lactic acid; L phenyllactic acid; 1'-(8alpha-FAD)-L-histidine; L-2'.4',5'-topaquinone; L-3',4'-dihydroxyphenylalanine; L-3'.4'.5'-trihydroxyphenylalanine; L-4'-

30   bromophenylalanine; L-6'-bromotryptophan; L-alanine amide; L-alanyl imidazolinone glycine; L-allysine; L-arginine amide; L-asparagine amide; L-aspartic 4-phosphoric anhydride; L-aspartic acid 1-amide; L-beta-methylthioaspartic acid; L-bromohistidine; L-citrulline; L-cysteine amide; L-cysteine glutathione disulfide; L-cysteine methyl disulfide;

L-cysteine methyl ester; L-cysteine oxazolecarboxylic acid; L-cysteine

oxazolinecarboxylic acid; L-cysteine persulfide; L-cysteine sulfenic acid; L-cysteine

sulfinic acid; L-cysteine thiazolecarboxylic acid; L-cysteinyl homocitryl molybdenum-

heptairon-nonasulfide; L-cysteinyl imidazolinone glycine; L-cysteinyl molybdopterin; L-

5   cysteinyl molybdopterin guanine dinucleotide; L-cystine; L-erythro-beta-

hydroxyasparagine; L-erythro-beta-hydroxyaspartic acid; L-gamma-carboxyglutamic acid;

L-glutamic acid 1-amide; L-glutamic acid 5-methyl ester; L-glutamine amide; L-glutamyl

5-glycerylphosphorylethanolamine; L-histidine amide; L-isoglutamyl-polyglutamic acid;

L-isoglutamyl-polyglycine; L-isoleucine amide; L-lanthionine; L-leucine amide; L-lysine

10   amide; L-lysine thiazolecarboxylic acid; L-lysinoalanine; L-methionine amide; L-

methionine sulfone; L-phenylalanine thiazolecarboxylic acid; L-phenylalanine amide; L-

proline amide; L-selenocysteine; L-selenocysteinyl molybdopterin guanine dinucleotide;

L-serine amide; L-serine thiazolecarboxylic acid; L-seryl imidazolinone glycine; L-T-

bromophenylalanine; L-T-bromophenylalanine; L-threonine amide; L-thyroxine; L-

15   tryptophan amide; L-tryptophyl quinone; L-tyrosine amide; L-valine amide; meso-

lanthionine; N-(L-glutamyl)-L-tyrosine; N-(L-isoaspartyl)-glycine; N-(L-isoaspartyl)-L-

cysteine; N,N,N-trimethyl-L-alanine; N,N-dimethyl-L-proline; N2-acetyl-L-lysine; N2-

succinyl-L-tryptophan; N4-(ADP-ribosyl)-L-asparagine; N4-glycosyl-L-asparagine; N4-

hydroxymethyl-L-asparagine; N4-methyl-L-asparagine; N5-methyl-L-glutamine; N6- 1 -

20   carboxyethyl-L-lysine; N6-(4-amino hydroxybutyl)-L-lysine; N6-(L-isoglutamyl)-L-

lysine; N6-(phospho-5'-adenosine)-L-lysine; N6-(phospho-5'-guanosine)-L-tysine;

N6,N6,N6-trimethyl-L-lysine; N6,N6-dimethyl-L-lysine; N6-acetyl-L-lysine; N6-biotinyl-

L-lysine; N6-carboxy-L-lysine; N6-formyl-L-lysine; N6-glycyl-L-lysine; N6-lipoyl-L-

lysine; N6-methyl-L-lysine; N6-methyl-N6-poly(N-methyl-propylamine)-L-lysine; N6-

25   mureinyl-L-lysine; N6-myristoyl-L-lysine; N6-palmitoyl-L-lysine; N6-pyridoxal

phosphate-L-lysine; N6-pyruvic acid 2-iminyl-L-lysine; N6-retinal-L-lysine; N-

acetylglycine; N-acetyl-L-glutamine; N-acetyl-L-alanine; N-acetyl-L-aspartic acid; N-

acetyl-L-cysteine; N-acetyl-L-glutamic acid; N-acetyl-L-isoleucine; N-acetyl-L-

methionine; N-acetyl-L-proline; N-acetyl-L-serine; N-acetyl-L-threonine; N-acetyl-L-

30   tyrosine; N-acetyl-L-valine; N-alanyl-glycosylphosphatidylinositolethanolamine; N-

asparaginyl-glycosylphosphatidylinositolethanolamine; N-aspartyl-

glycosylphosphatidylinositolethanolamine; N-formylglycine; N-formyl-L-methionine; N-

glycyl-glycosylphosphatidylinositolethanolamine; N-L-glutamyl-poly-L-glutamic acid; N-

methylglycine; N-methyl-L-alanine; N-methyl-L-methionine; N-methyl-L-phenylalanine; N-myristoyl-glycine; N-palmitoyl-L-cysteine; N-pyruvic acid 2-iminyl-L-cysteine; N-pyruvic acid 2-iminyl-L-valine; N-seryl-glycosylphosphatidylinositolethanolamine; N-seryl-glycosyOSPhingolipidinositolethanolamine; O-(ADP-ribosyl)-L-serine; O-(phospho-

5   5'-adenosine)-L-threonine; O-(phospho-5'-DNA)-L-serine; O-(phospho-5'-DNA)-L-threonine; O-(phospho-5'rRNA)-L-serine; O-(phosphoribosyl dephospho-coenzyme A)-L-serine; O-(sn-1-glycerophosphoryl)-L-serine; O4'-(8alpha-FAD)-L-tyrosine; O4'-(phospho-5'-adenosine)-L-tyrosine; O4'-(phospho-5'-DNA)-L-tyrosine; O4'-(phospho-5'-RNA)-L-tyrosine; O4'-(phospho-5'-uridine)-L-tyrosine; O4-glycosyl-L-hydroxyproline; O4'-

10   glycosyl-L-tyrosine; O4'-sulfo-L-tyrosine; O5-glycosyl-L-hydroxylysine; O-glycosyl-L-serine; O-glycosyl-L-threonine; omega-N-(ADP-ribosyl)-L-arginine; omega-N-omega-N'-dimethyl-L-arginine; omega-N-methyl-L-arginine; omega-N-omega-N-dimethyl-L-arginine; omega-N-phospho-L-arginine; O'octanoyl-L-serine; O-palmitoyl-L-serine; O-palmitoyl-L-threonine; O-phospho-L-serine; O-phospho-L-threonine; O-

15   phosphopantetheine-L-serine; phycoerythrobilin-bis-L-cysteine; phycourobilin-bis-L-cysteine; pyrroloquinoline quinone; pyruvic acid; S hydroxycinnamyl-L-cysteine; S-(2-aminovinyl) methyl-D-eysteine; S-(2-aminovinyl)-D-cysteine; S-(6-FW-L-cysteine; S-(8alpha-FAD)-L-cysteine; S-(ADP-ribosyl)-L-cysteine; S-(L-isoglutamyl)-L-cysteine; S-12-hydroxyfarnesyl-L-cysteine; S-acetyl-L-cysteine; S-diacylglycerol-L-cysteine; S-

20   diphytanylglycerot diether-L-cysteine; S-farnesyl-L-cysteine; S-geranylgeranyl-L-cysteine; S-glycosyl-L-cysteine; S-glycyl-L-cysteine; S-methyl-L-cysteine; S-nitrosyl-L-cysteine; S-palmitoyl-L-cysteine; S-phospho-L-cysteine; S-phycobiliviolin-L-cysteine; S-phycocyanobilin-L-cysteine; S-phycoerythrobilin-L-cysteine; S-phytochromobilin-L-cysteine; S-selenyl-L-cysteine; S-sulfo-L-cysteine; tetrakis-L-cysteinyl diiron disulfide;

25   tetrakis-L-cysteinyl iron; tetrakis-L-cysteinyl tetrairon tetrasulfide; trans-2,3-cis 4-dihydroxy-L-proline; tris-L-cysteinyl triiron tetrasulfide; tris-L-cysteinyl triiron trisulfide; tris-L-cysteinyl-L-aspartato tetrairon tetrasulfide; tris-L-cysteinyl-L-cysteine persulfido-bis-L-glutamato-L-histidino tetrairon disulfide trioxide; tris-L-cysteinyl-L-N3'-histidino tetrairon tetrasulfide; tris-L-cysteinyl-L-NI'-histidino tetrairon tetrasulfide; and tris-L-

30   cysteinyl-L-serinyl tetrairon tetrasulfide.

Additional examples of PTMs may be found in web sites such as the Delta Mass database based on Krishna, R. G. and F. Wold (1998). Posttranslational Modifications. Proteins - Analysis and Design. R. H. Angeletti. San Diego, Academic Press. 1: 121-206. ;

Methods in Enzymology, 193, J.A. McClosky (ed) (1990), pages 647-660; Methods in Protein Sequence Analysis edited by Kazutomo Imahori and Fumio Sakiyama, Plenum Press, (1993) "Post-translational modifications of proteins" R.G. Krishna and F. Wold pages 167-172; "GlycoSuiteDB: a new curated relational database of glycoprotein glycan

5   structures and their biological sources" Cooper et al. Nucleic Acids Res. 29; 332-335 (2001) "O-GLYCBASE version 4.0: a revised database of O-glycosylated proteins" Gupta et al. Nucleic Acids Research, 27: 370-372 (1999); and "PhosphoBase, a database of phosphorylation sites: release 2.0.", Kreegipuu et al.Nucleic Acids Res 27(1):237-239 (1999) see also,WO 02/21139A2, the disclosure of which is incorporated herein by

10   reference in its entirety.

Tumorigenesis is often accompanied by alterations in the post-translational modifications of proteins. Thus, in another embodiment, the invention provides polypeptides from cancerous cells or tissues that have altered post-translational modifications compared to the post-translational modifications of polypeptides from

15   normal cells or tissues. A number of altered post-translational modifications are known. One common alteration is a change in phosphorylation state, wherein the polypeptide from the cancerous cell or tissue is hyperphosphorylated or hypophosphorylated compared to the polypeptide from a normal tissue, or wherein the polypeptide is phosphorylated on different residues than the polypeptide from a normal cell. Another common alteration is

20   a change in glycosylation state, wherein the polypeptide from the cancerous cell or tissue has more or less glycosylation than the polypeptide from a normal tissue, and/or wherein the polypeptide from the cancerous cell or tissue has a different type of glycosylation than the polypeptide from a noncancerous cell or tissue. Changes in glycosylation may be critical because carbohydrate-protein and carbohydrate-carbohydrate interactions are

25   important in cancer cell progression, dissemination and invasion. See, e.g., Barchi, *Curr. Pharm. Des.* 6: 485-501 (2000), Verma, *Cancer Biochem. Biophys.* 14: 151-162 (1994) and Dennis et al., *Bioessays* 5: 412-421 (1999).

Another post-translational modification that may be altered in cancer cells is prenylation. Prenylation is the covalent attachment of a hydrophobic prenyl group (either

30   farnesyl or geranylgeranyl) to a polypeptide. Prenylation is required for localizing a protein to a cell membrane and is often required for polypeptide function. For instance, the Ras superfamily of GTPase signalling proteins must be prenylated for function in a

cell. See, e.g., Prendergast et al., *Semin. Cancer Biol.* 10: 443-452 (2000) and Khwaja et al., *Lancet* 355: 741-744 (2000).

Other post-translation modifications that may be altered in cancer cells include, without limitation, polypeptide methylation, acetylation, arginylation or racemization of

5   amino acid residues. In these cases, the polypeptide from the cancerous cell may exhibit either increased or decreased amounts of the post-translational modification compared to the corresponding polypeptides from noncancerous cells.

Other polypeptide alterations in cancer cells include abnormal polypeptide cleavage of proteins and aberrant protein-protein interactions. Abnormal polypeptide

10   cleavage may be cleavage of a polypeptide in a cancerous cell that does not usually occur in a normal cell, or a lack of cleavage in a cancerous cell, wherein the polypeptide is cleaved in a normal cell. Aberrant protein-protein interactions may be either covalent cross-linking or non-covalent binding between proteins that do not normally bind to each other. Alternatively, in a cancerous cell, a protein may fail to bind to another protein to

15   which it is bound in a noncancerous cell. Alterations in cleavage or in protein-protein interactions may be due to over- or underproduction of a polypeptide in a cancerous cell compared to that in a normal cell, or may be due to alterations in post-translational modifications (see above) of one or more proteins in the cancerous cell. See, e.g., Henschen-Edman, *Ann. N.Y. Acad. Sci.* 936: 580-593 (2001).

20   Alterations in polypeptide post-translational modifications, as well as changes in polypeptide cleavage and protein-protein interactions, may be determined by any method known in the art. For instance, alterations in phosphorylation may be determined by using anti-phosphoserine, anti-phosphothreonine or anti-phosphotyrosine antibodies or by amino acid analysis. Glycosylation alterations may be determined using antibodies specific for

25   different sugar residues, by carbohydrate sequencing, or by alterations in the size of the glycoprotein, which can be determined by, e.g., SDS polyacrylamide gel electrophoresis (PAGE). Other alterations of post-translational modifications, such as prenylation, racemization, methylation, acetylation and arginylation, may be determined by chemical analysis, protein sequencing, amino acid analysis, or by using antibodies specific for the

30   particular post-translational modifications. Changes in protein-protein interactions and in polypeptide cleavage may be analyzed by any method known in the art including, without limitation, non-denaturing PAGE (for non-covalent protein-protein interactions), SDS

PAGE (for covalent protein-protein interactions and protein cleavage), chemical cleavage, protein sequencing or immunoassays.

In another embodiment, the invention provides polypeptides that have been post-translationally modified. In one embodiment, polypeptides may be modified

5   enzymatically or chemically, by addition or removal of a post-translational modification. For example, a polypeptide may be glycosylated or deglycosylated enzymatically. Similarly, polypeptides may be phosphorylated using a purified kinase, such as a MAP kinase (e.g, p38, ERK, or JNK) or a tyrosine kinase (e.g., Src or erbB2). A polypeptide may also be modified through synthetic chemistry. Alternatively, one may isolate the

10 polypeptide of interest from a cell or tissue that expresses the polypeptide with the desired post-translational modification. In another embodiment, a nucleic acid molecule encoding the polypeptide of interest is introduced into a host cell that is capable of post-translationally modifying the encoded polypeptide in the desired fashion. If the polypeptide does not contain a motif for a desired post-translational modification, one may

15 alter the post-translational modification by mutating the nucleic acid sequence of a nucleic acid molecule encoding the polypeptide so that it contains a site for the desired post-translational modification. Amino acid sequences that may be post-translationally modified are known in the art. See, e.g., the programs described above on the website www.expasy.org. The nucleic acid molecule may also be introduced into a host cell that is

20 capable of post-translationally modifying the encoded polypeptide. Similarly, one may delete sites that are post-translationally modified by either mutating the nucleic acid sequence so that the encoded polypeptide does not contain the post-translational modification motif, or by introducing the native nucleic acid molecule into a host cell that is not capable of post-translationally modifying the encoded polypeptide.

25   It will be appreciated, as is well known and as noted above, that polypeptides are not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslation events, including natural processing event and events brought about by human manipulation which do not occur naturally. Circular, branched and branched

30 circular polypeptides may be synthesized by non-translation natural process and by entirely synthetic methods, as well. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl group in a polypeptide, or both, by a

covalent modification, is common in naturally occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli*, prior to proteolytic processing, almost invariably will be N-formylmethionine.

5   Useful post-synthetic (and post-translational) modifications include conjugation to detectable labels, such as fluorophores. A wide variety of amine-reactive and thiol-reactive fluorophore derivatives have been synthesized that react under nondenaturing conditions with N-terminal amino groups and epsilon amino groups of lysine residues, on the one hand, and with free thiol groups of cysteine residues, on the other.

10   Kits are available commercially that permit conjugation of proteins to a variety of amine-reactive or thiol-reactive fluorophores: Molecular Probes, Inc. (Eugene, OR, USA), *e.g.*, offers kits for conjugating proteins to Alexa Fluor 350, Alexa Fluor 430, Fluorescein-EX, Alexa Fluor 488, Oregon Green 488, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor 546, Alexa Fluor 568, Alexa Fluor 594, and Texas Red-X.

15   A wide variety of other amine-reactive and thiol-reactive fluorophores are available commercially (Molecular Probes, Inc., Eugene, OR, USA), including Alexa Fluor® 350, Alexa Fluor® 488, Alexa Fluor® 532, Alexa Fluor® 546, Alexa Fluor® 568, Alexa Fluor® 594, Alexa Fluor® 647 (monoclonal antibody labeling kits available from Molecular Probes, Inc., Eugene, OR, USA), BODIPY dyes, such as BODIPY 493/503,

20   BODIPY FL, BODIPY R6G, BODIPY 530/550, BODIPY TMR, BODIPY 558/568, BODIPY 558/568, BODIPY 564/570, BODIPY 576/589, BODIPY 581/591, BODIPY TR, BODIPY 630/650, BODIPY 650/665, Cascade Blue, Cascade Yellow, Dansyl, lissamine rhodamine B, Marina Blue, Oregon Green 488, Oregon Green 514, Pacific Blue, rhodamine 6G, rhodamine green, rhodamine red, tetramethylrhodamine, Texas Red

25   (available from Molecular Probes, Inc., Eugene, OR, USA).

The polypeptides of the present invention can also be conjugated to fluorophores, other proteins, and other macromolecules, using bifunctional linking reagents. Common homobifunctional reagents include, *e.g.*, APG, AEDP, BASED, BMB, BMDB, BMH, BMOE, BM[PEO]3, BM[PEO]4, BS3, BSOCOES, DFDNB, DMA, DMP, DMS, DPDPB,

30   DSG, DSP (Lomant's Reagent), DSS, DST, DTBP, DTME, DTSSP, EGS, HBVS, Sulfo-BSOCOES, Sulfo-DST, Sulfo-EGS (all available from Pierce, Rockford, IL, USA); common heterobifunctional cross-linkers include ABH, AMAS, ANB-NOS, APDP, ASBA, BMPA, BMPH, BMPS, EDC, EMCA, EMCH, EMCS, KMUA, KMUH, GMBS,

LC-SMCC, LC-SPDP, MBS, M2C2H, MPBH, MSA, NHS-ASA, PDPH, PMPI, SADP,

SAED, SAND, SANPAH, SASD, SATP, SBAP, SFAD, SIA, SIAB, SMCC, SMPB,

SMPH, SMPT, SPDP, Sulfo-EMCS, Sulfo-GMBS, Sulfo-HSAB, Sulfo-KMUS,

Sulfo-LC-SPDP, Sulfo-MBS, Sulfo-NHS-LC-ASA, Sulfo-SADP, Sulfo-SANPAH,

5   Sulfo-SIAB, Sulfo-SMCC, Sulfo-SMPB, Sulfo-LC-SMPT, SVSB, TFCS (all available

Pierce, Rockford, IL, USA).

Polypeptides of the present invention, including full length polypeptides,

fragments and fusion proteins, can be conjugated, using such cross-linking reagents, to

fluorophores that are not amine- or thiol-reactive. Other labels that usefully can be

10   conjugated to polypeptides of the present invention include radioactive labels,

echosonographic contrast reagents, and MRI contrast agents.

Polypeptides of the present invention, including full length polypeptide, fragments

and fusion proteins, can also usefully be conjugated using cross-linking agents to carrier

proteins, such as KLH, bovine thyroglobulin, and even bovine serum albumin (BSA), to

15   increase immunogenicity for raising anti-OSP antibodies.

Polypeptides of the present invention, including full length polypeptide, fragments

and fusion proteins, can also usefully be conjugated to polyethylene glycol (PEG);

PEGylation increases the serum half life of proteins administered intravenously for

replacement therapy. Delgado *et al.*, *Crit. Rev. Ther. Drug Carrier Syst.* 9(3-4): 249-304

20   (1992); Scott *et al.*, *Curr. Pharm. Des.* 4(6): 423-38 (1998); DeSantis *et al.*, *Curr. Opin.*

Biotechnol. 10(4): 324-30 (1999). PEG monomers can be attached to the protein directly

or through a linker, with PEGylation using PEG monomers activated with tresyl chloride

(2,2,2-trifluoroethanesulphonyl chloride) permitting direct attachment under mild·

conditions.

25   Polypeptides of the present invention are also inclusive of analogs of a polypeptide

encoded by a nucleic acid molecule according to the instant invention. In a preferred

embodiment, this polypeptide is a OSP. In a more preferred embodiment, this polypeptide

is derived from a polypeptide having part or all of the amino acid sequence of SEQ ID

NO: 100-192. Also preferred is an analog polypeptide comprising one or more

30   substitutions of non-natural amino acids or non-native inter-residue bonds compared to the

naturally occurring polypeptide. In one embodiment, the analog is structurally similar to a

OSP, but one or more peptide linkages is replaced by a linkage selected from the group

consisting of --CH_2NH--, --CH_2S--, --CH_2-CH_2--, --CH=CH--(cis and trans), --$COCH_2$--, ·

--CH(OH)CH$_2$-- and --CH$_2$SO--. In another embodiment, the analog comprises

substitution of one or more amino acids of a OSP with a D-amino acid of the same type or

other non-natural amino acid in order to generate more stable peptides. D-amino acids can

readily be incorporated during chemical peptide synthesis: peptides assembled from

5   D-amino acids are more resistant to proteolytic attack; incorporation of D-amino acids can

also be used to confer specific three-dimensional conformations on the peptide. Other

amino acid analogues commonly added during chemical synthesis include ornithine,

norleucine, phosphorylated amino acids (typically phosphoserine, phosphothreonine,

phosphotyrosine), L-malonyltyrosine, a non-hydrolyzable analog of phosphotyrosine (see,

10   e.g., Kole et al., Biochem. Biophys. Res. Com. 209: 817-821 (1995)), and various

halogenated phenylalanine derivatives.

       Non-natural amino acids can be incorporated during solid phase chemical synthesis

       or by recombinant techniques, although the former is typically more common. Solid

       phase chemical synthesis of peptides is well established in the art. Procedures are

15   described, inter alia, in Chan et al. (eds.), Fmoc Solid Phase Peptide Synthesis: A-

Practical Approach (Practical Approach Series), Oxford Univ. Press (March 2000); Jones,

Amino Acid and Peptide Synthesis (Oxford Chemistry Primers, No 7), Oxford Univ. Press

(1992); and Bodanszky, Principles of Peptide Synthesis (Springer Laboratory), Springer

Verlag (1993).

20   Amino acid analogues having detectable labels are also usefully incorporated

during synthesis to provide derivatives and analogs. Biotin, for example can be added

using biotinoyl--(9-fluorenylmethoxycarbonyl)-L-lysine (FMOC biocytin) (Molecular

Probes, Eugene, OR, USA). Biotin can also be added enzymatically by incorporation into

a fusion protein of a E. coli BirA substrate peptide. The FMOC and tBOC derivatives of

25   dabcyl-L-lysine (Molecular Probes, Inc., Eugene, OR, USA) can be used to incorporate

the dabcyl chromophore at selected sites in the peptide sequence during synthesis. The

aminonaphthalene derivative EDANS, the most common fluorophore for pairing with the

dabcyl quencher in fluorescence resonance energy transfer (FRET) systems, can be

introduced during automated synthesis of peptides by using EDANS--FMOC-L-glutamic

30   acid or the corresponding tBOC derivative (both from Molecular Probes, Inc., Eugene,

OR, USA). Tetramethylrhodamine fluorophores can be incorporated during automated

FMOC synthesis of peptides using (FMOC)--TMR-L-lysine (Molecular Probes, Inc.

Eugene, OR, USA).

Other useful amino acid analogues that can be incorporated during chemical synthesis include aspartic acid, glutamic acid, lysine, and tyrosine analogues having allyl side-chain protection (Applied Biosystems, Inc., Foster City, CA, USA); the allyl side chain permits synthesis of cyclic, branched-chain, sulfonated, glycosylated, and

5   phosphorylated peptides.

A large number of other FMOC-protected non-natural amino acid analogues capable of incorporation during chemical synthesis are available commercially, including, e.g., Fmoc-2-aminobicyclo[2.2.1]heptane-2-carboxylic acid, Fmoc-3-endo-aminobicyclo[2.2.1]heptane-2-endo-carboxylic acid, Fmoc-3-exo-

10   aminobicyclo[2.2.1]heptane-2-exo-carboxylic acid, Fmoc-3-endo-amino-bicyclo[2.2.1]hept-5-ene-2-endo-carboxylic acid, Fmoc-3-exo-amino-bicyclo[2.2.1]hept-5-ene-2-exo-carboxylic acid, Fmoc-cis-2-amino-1-cyclohexanecarboxylic acid, Fmoc-trans-2-amino-1-cyclohexanecarboxylic acid, Fmoc-1-amino-1-cyclopentanecarboxylic acid, Fmoc-cis-2-amino-1-cyclopentanecarboxylic acid, Fmoc-1-amino-1-

15   cyclopropanecarboxylic acid, Fmoc-D-2-amino-4-(ethylthio)butyric acid, Fmoc-L-2-amino-4-(ethylthio)butyric acid, Fmoc-L-buthionine, Fmoc-S-methyl-L-Cysteine, Fmoc-2-aminobenzoic acid (anthranillic acid), Fmoc-3-aminobenzoic acid, Fmoc-4-aminobenzoic acid, Fmoc-2-aminobenzophenone-2'-carboxylic acid, Fmoc-N-(4-aminobenzoyl)-β-alanine, Fmoc-2-amino-4,5-dimethoxybenzoic acid, Fmoc-4-

20   aminohippuric acid, Fmoc-2-amino-3-hydroxybenzoic acid, Fmoc-2-amino-5-hydroxybenzoic acid, Fmoc-3-amino-4-hydroxybenzoic acid, Fmoc-4-amino-3-hydroxybenzoic acid, Fmoc-4-amino-2-hydroxybenzoic acid, Fmoc-5-amino-2-hydroxybenzoic acid, Fmoc-2-amino-3-methoxybenzoic acid, Fmoc-4-amino-3-methoxybenzoic acid, Fmoc-2-amino-3-methylbenzoic acid, Fmoc-2-amino-5-

25   methylbenzoic acid, Fmoc-2-amino-6-methylbenzoic acid, Fmoc-3-amino-2-methylbenzoic acid, Fmoc-3-amino-4-methylbenzoic acid, Fmoc-4-amino-3-methylbenzoic acid, Fmoc-3-amino-2-naphtoic acid, Fmoc-D,L-3-amino-3-phenylpropionic acid, Fmoc-L-Methyldopa, Fmoc-2-amino-4,6-dimethyl-3-pyridinecarboxylic acid, Fmoc-D,L-amino-2-thiophenacetic acid, Fmoc-4-

30   (carboxymethyl)piperazine, Fmoc-4-carboxypiperazine, Fmoc-4-(carboxymethyl)homopiperazine, Fmoc-4-phenyl-4-piperidinecarboxylic acid, Fmoc-L-1,2,3,4-tetrahydronorharman-3-carboxylic acid, Fmoc-L-thiazolidine-4-carboxylic acid, all available from The Peptide Laboratory (Richmond, CA, USA).

Non-natural residues can also be added biosynthetically by engineering a suppressor tRNA, typically one that recognizes the UAG stop codon, by chemical aminoacylation with the desired unnatural amino acid. Conventional site-directed mutagenesis is used to introduce the chosen stop codon UAG at the site of interest in the

5 protein gene. When the acylated suppressor tRNA and the mutant gene are combined in an *in vitro* transcription/translation system, the unnatural amino acid is incorporated in response to the UAG codon to give a protein containing that amino acid at the specified position. Liu *et al.*, *Proc. Natl Acad. Sci. USA* 96(9): 4780-5 (1999); Wang *et al.*, *Science* 292(5516): 498-500 (2001).

10 *Fusion Proteins*

Another aspect of the present invention relates to the fusion of a polypeptide of the present invention to heterologous polypeptides. In a preferred embodiment, the polypeptide of the present invention is a OSP. In a more preferred embodiment, the polypeptide of the present invention that is fused to a heterologous polypeptide comprises

15 part or all of the amino acid sequence of SEQ ID NO: 100-192, or is a mutein, homologous polypeptide, analog or derivative thereof. In an even more preferred embodiment, the fusion protein is encoded by a nucleic acid molecule comprising all or part of the nucleic acid sequence of SEQ ID NO: 1-99, or comprises all or part of a nucleic acid sequence that selectively hybridizes or is homologous to a nucleic acid molecule

20 comprising a nucleic acid sequence of SEQ ID NO: 1-99.

The fusion proteins of the present invention will include at least one fragment of a polypeptide of the present invention, which fragment is at least 6, typically at least 8, often at least 15, and usefully at least 16, 17, 18, 19, or 20 amino acids long. The fragment of the polypeptide of the present to be included in the fusion can usefully be at least 25

25 amino acids long, at least 50 amino acids long, and can be at least 75, 100, or even 150 amino acids long. Fusions that include the entirety of a polypeptide of the present invention have particular utility.

The heterologous polypeptide included within the fusion protein of the present invention is at least 6 amino acids in length, often at least 8 amino acids in length, and

30 preferably at least 15, 20, or 25 amino acids in length. Fusions that include larger polypeptides, such as the IgG Fc region, and even entire proteins (such as GFP chromophore-containing proteins) are particularly useful.

As described above in the description of vectors and expression vectors of the present invention, which discussion is incorporated here by reference in its entirety, heterologous polypeptides to be included in the fusion proteins of the present invention can usefully include those designed to facilitate purification and/or visualization of

5   recombinantly-expressed proteins. *See, e.g.,* Ausubel, Chapter 16, (1992), *supra.* Although purification tags can also be incorporated into fusions that are chemically synthesized, chemical synthesis typically provides sufficient purity that further purification by HPLC suffices; however, visualization tags as above described retain their utility even when the protein is produced by chemical synthesis, and when so included

10   render the fusion proteins of the present invention useful as directly detectable markers of the presence of a polypeptide of the invention.

As also discussed above, heterologous polypeptides to be included in the fusion proteins of the present invention can usefully include those that facilitate secretion of recombinantly expressed proteins into the periplasmic space or extracellular milieu for

15   prokaryotic hosts or into the culture medium for eukaryotic cells through incorporation of secretion signals and/or leader sequences. For example, a His[6] tagged protein can be purified on a Ni affinity column and a GST fusion protein can be purified on a glutathione affinity column. Similarly, a fusion protein comprising the Fc domain of IgG can be purified on a Protein A or Protein G column and a fusion protein comprising an epitope

20   tag such as myc can be purified using an immunoaffinity column containing an anti-c-myc antibody. It is preferable that the epitope tag be separated from the protein encoded by the essential gene by an enzymatic cleavage site that can be cleaved after purification. See also the discussion of nucleic acid molecules encoding fusion proteins that may be expressed on the surface of a cell.

25   Other useful fusion proteins of the present invention include those that permit use of the polypeptide of the present invention as bait in a yeast two-hybrid system. *See* Bartel *et al.* (eds.), The Yeast Two-Hybrid System, Oxford University Press (1997); Zhu *et al.*, Yeast Hybrid Technologies, Eaton Publishing (2000); Fields *et al.*, *Trends Genet.* 10(8): 286-92 (1994); Mendelsohn *et al.*, *Curr. Opin. Biotechnol.* 5(5): 482-6 (1994);

30   Luban *et al.*, *Curr. Opin. Biotechnol.* 6(1): 59-64 (1995); Allen *et al.*, *Trends Biochem. Sci.* 20(12): 511-6 (1995); Drees, *Curr. Opin. Chem. Biol.* 3(1): 64-70 (1999); Topcu *et al.*, *Pharm. Res.* 17(9): 1049-55 (2000); Fashena *et al.*, *Gene* 250(1-2): 1-14 (2000); Colas *et al.*, *Nature* 380, 548-550 (1996); Norman, T. *et al.*, *Science* 285, 591-595 (1999);

Fabbrizio *et al.*, *Oncogene* 18, 4357-4363 (1999); Xu *et al.*, *Proc Natl Acad Sci U S A.* 94, 12473-12478 (1997); Yang, *et al.*, *Nuc. Acids Res.* 23, 1152-1156 (1995); Kolonin *et al.*, *Proc Natl Acad Sci U S A* 95, 14266-14271 (1998); Cohen *et al.*, *Proc Natl Acad Sci U S A* 95, 14272-14277 (1998); Uetz, *et al. Nature* 403, 623-627(2000); Ito, *et al.*, *Proc Natl Acad Sci U S A* 98, 4569-4574 (2001). Typically, such fusion is to either *E. coli* LexA or yeast GAL4 DNA binding domains. Related bait plasmids are available that express the bait fused to a nuclear localization signal.

Other useful fusion proteins include those that permit display of the encoded polypeptide on the surface of a phage or cell, fusions to intrinsically fluorescent proteins, such as green fluorescent protein (GFP), and fusions to the IgG Fc region, as described above.

The polypeptides of the present invention can also usefully be fused to protein toxins, such as Pseudomonas exotoxin A, diphtheria toxin, shiga toxin A, anthrax toxin lethal factor, ricin, in order to effect ablation of cells that bind or take up the proteins of the present invention.

Fusion partners include, *inter alia, myc*, hemagglutinin (HA), GST, immunoglobulins, β-galactosidase, biotin trpE, protein A, β-lactamase, α-amylase, maltose binding protein, alcohol dehydrogenase, polyhistidine (for example, six histidine at the amino and/or carboxyl terminus of the polypeptide), lacZ, green fluorescent protein (GFP), yeast α mating factor, GAL4 transcription activation or DNA binding domain, luciferase, and serum proteins such as ovalbumin, albumin and the constant domain of IgG. *See, e.g.*, Ausubel (1992), *supra* and Ausubel (1999), *supra*. Fusion proteins may also contain sites for specific enzymatic cleavage, such as a site that is recognized by enzymes such as Factor XIII, trypsin, pepsin, or any other enzyme known in the art. Fusion proteins will typically be made by either recombinant nucleic acid methods, as described above, chemically synthesized using techniques well known in the art (*e.g.*, a Merrifield synthesis), or produced by chemical cross-linking.

Another advantage of fusion proteins is that the epitope tag can be used to bind the fusion protein to a plate or column through an affinity linkage for screening binding proteins or other molecules that bind to the OSP.

As further described below, the polypeptides of the present invention can readily be used as specific immunogens to raise antibodies that specifically recognize polypeptides of the present invention including OSPs and their allelic variants and

homologues. The antibodies, in turn, can be used, *inter alia*, specifically to assay for the polypeptides of the present invention, particularly OSPs, *e.g.* by ELISA for detection of protein fluid samples, such as serum, by immunohistochemistry or laser scanning cytometry, for detection of protein in tissue samples, or by flow cytometry, for detection

5   of intracellular protein in cell suspensions, for specific antibody-mediated isolation and/or purification of OSPs, as for example by immunoprecipitation, and for use as specific agonists or antagonists of OSPs.

One may determine whether polypeptides of the present invention including OSPs, muteins, homologous proteins or allelic variants or fusion proteins of the present invention

10   are functional by methods known in the art. For instance, residues that are tolerant of change while retaining function can be identified by altering the polypeptide at known residues using methods known in the art, such as alanine scanning mutagenesis, Cunningham *et al.*, *Science* 244(4908): 1081-5 (1989); transposon linker scanning mutagenesis, Chen *et al.*, *Gene* 263(1-2): 39-48 (2001); combinations of homolog- and

15   alanine-scanning mutagenesis, Jin *et al.*, *J. Mol. Biol.* 226(3): 851-65 (1992); combinatorial alanine scanning, Weiss *et al.*, *Proc. Natl. Acad. Sci USA* 97(16): 8950-4 (2000), followed by functional assay. Transposon linker scanning kits are available commercially (New England Biolabs, Beverly, MA, USA, catalog. no. E7-102S; EZ::TN™ In-Frame Linker Insertion Kit, catalogue no. EZI04KN, (Epicentre

20   Technologies Corporation, Madison, WI, USA).

Purification of the polypeptides or fusion proteins of the present invention is well known and within the skill of one having ordinary skill in the art. *See, e.g.*, Scopes, Protein Purification, 2d ed. (1987).. Purification of recombinantly expressed polypeptides is described above. Purification of chemically-synthesized peptides can readily be

25   effected, *e.g.*, by HPLC.

Accordingly, it is an aspect of the present invention to provide the isolated polypeptides or fusion proteins of the present invention in pure or substantially pure form in the presence of absence of a stabilizing agent. Stabilizing agents include both proteinaceous and non-proteinaceous material and are well known in the art. Stabilizing

30   agents, such as albumin and polyethylene glycol (PEG) are known and are commercially available.

Although high levels of purity are preferred when the isolated polypeptide or fusion protein of the present invention are used as therapeutic agents, such as in vaccines

and replacement therapy, the isolated polypeptides of the present invention are also useful at lower purity. For example, partially purified polypeptides of the present invention can be used as immunogens to raise antibodies in laboratory animals.

In a preferred embodiment, the purified and substantially purified polypeptides of

5   the present invention are in compositions that lack detectable ampholytes, acrylamide monomers, bis-acrylamide monomers, and polyacrylamide.

The polypeptides or fusion proteins of the present invention can usefully be attached to a substrate. The substrate can be porous or solid, planar or non-planar; the bond can be covalent or noncovalent. For example, the peptides of the invention may be

10   stabilized by covalent linkage to albumin. See, U.S. Patent No. 5,876,969, the contents of which are hereby incorporated in its entirety.

For example, the polypeptides or fusion proteins of the present invention can usefully be bound to a porous substrate, commonly a membrane, typically comprising nitrocellulose, polyvinylidene fluoride (PVDF), or cationically derivatized, hydrophilic

15   PVDF; so bound, the polypeptides or fusion proteins of the present invention can be used to detect and quantify antibodies, *e.g.* in serum, that bind specifically to the immobilized polypeptide or fusion protein of the present invention.

As another example, the polypeptides or fusion proteins of the present invention can usefully be bound to a substantially nonporous substrate, such as plastic, to detect and

20   quantify antibodies, *e.g.* in serum, that bind specifically to the immobilized protein of the present invention. Such plastics include polymethylacrylic, polyethylene, polypropylene, polyacrylate, polymethylmethacrylate, polyvinylchloride, polytetrafluoroethylene, polystyrene, polycarbonate, polyacetal, polysulfone, celluloseacetate, cellulosenitrate, nitrocellulose, or mixtures thereof; when the assay is performed in a standard microtiter

25   dish, the plastic is typically polystyrene.

The polypeptides and fusion proteins of the present invention can also be attached to a substrate suitable for use as a surface enhanced laser desorption ionization source; so attached, the polypeptide or fusion protein of the present invention is useful for binding and then detecting secondary proteins that bind with sufficient affinity or avidity to the

30   surface-bound polypeptide or fusion protein to indicate biologic interaction there between. The polypeptides or fusion proteins of the present invention can also be attached to a substrate suitable for use in surface plasmon resonance detection; so attached, the polypeptide or fusion protein of the present invention is useful for binding and then

detecting secondary proteins that bind with sufficient affinity or avidity to the surface-bound polypeptide or fusion protein to indicate biological interaction there between.

<u>Antibodies</u>

In another aspect, the invention provides antibodies, including fragments and derivatives thereof, that bind specifically to polypeptides encoded by the nucleic acid molecules of the invention. In a preferred embodiment, the antibodies are specific for a polypeptide that is a OSP, or a fragment, mutein, derivative, analog or fusion protein thereof. In a more preferred embodiment, the antibodies are specific for a polypeptide that comprises SEQ ID NO: 100-192, or a fragment, mutein, derivative, analog or fusion protein thereof.

The antibodies of the present invention can be specific for linear epitopes, discontinuous epitopes, or conformational epitopes of such proteins or protein fragments, either as present on the protein in its native conformation or, in some cases, as present on the proteins as denatured, as, *e.g.*, by solubilization in SDS. New epitopes may be also due to a difference in post translational modifications (PTMs) in disease versus normal tissue. For example, a particular site on a OSP may be glycosylated in cancerous cells, but not glycosylated in normal cells or vis versa. In addition, alternative splice forms of a OSP may be indicative of cancer. Differential degradation of the C or N-terminus of a OSP may also be a marker or target for anticancer therapy. For example, an OSP may be N-terminal degraded in cancer cells exposing new epitopes to which antibodies may selectively bind for diagnostic or therapeutic uses.

As is well known in the art, the degree to which an antibody can discriminate as among molecular species in a mixture will depend, in part, upon the conformational relatedness of the species in the mixture; typically, the antibodies of the present invention will discriminate over adventitious binding to non-OSP polypeptides by at least two-fold, more typically by at least 5-fold, typically by more than 10-fold, 25-fold, 50-fold, 75-fold, and often by more than 100-fold, and on occasion by more than 500-fold or 1000-fold. When used to detect the proteins or protein fragments of the present invention, the antibody of the present invention is sufficiently specific when it can be used to determine the presence of the polypeptide of the present invention in samples derived from human ovarian.

Typically, the affinity or avidity of an antibody (or antibody multimer, as in the case of an IgM pentamer) of the present invention for a protein or protein fragment of the present invention will be at least about 1×10^{-6} molar (M), typically at least about 5×10^{-7} M, 1×10^{-7} M, with affinities and avidities of at least 1×10^{-8} M, 5×10^{-9} M, 1×10^{-10} M

5   and up to 1×10^{-13} M proving especially useful.

The antibodies of the present invention can be naturally occurring forms, such as IgG, IgM, IgD, IgE, IgY, and IgA, from any avian, reptilian, or mammalian species.

Human antibodies can, but will infrequently, be drawn directly from human donors or human cells. In such case, antibodies to the polypeptides of the present invention will

10   typically have resulted from fortuitous immunization, such as autoimmune immunization, with the polypeptide of the present invention. Such antibodies will typically, but will not invariably, be polyclonal. In addition, individual polyclonal antibodies may be isolated and cloned to generate monoclonals.

Human antibodies are more frequently obtained using transgenic animals that

15   express human immunoglobulin genes, which transgenic animals can be affirmatively immunized with the protein immunogen of the present invention. Human Ig-transgenic mice capable of producing human antibodies and methods of producing human antibodies therefrom upon specific immunization are described, *inter alia*, in U.S. Patent Nos. 6,162,963; 6,150,584; 6,114,598; 6,075,181; 5,939,598; 5,877,397; 5,874,299; 5,814,318;

20   5,789,650; 5,770,429; 5,661,016; 5,633,425; 5,625,126; 5,569,825; 5,545,807; 5,545,806, and 5,591,669, the disclosures of which are incorporated herein by reference in their entireties. Such antibodies are typically monoclonal, and are typically produced using techniques developed for production of murine antibodies.

Human antibodies are particularly useful, and often preferred, when the antibodies

25   of the present invention are to be administered to human beings as *in vivo* diagnostic or therapeutic agents, since recipient immune response to the administered antibody will often be substantially less than that occasioned by administration of an antibody derived from another species, such as mouse.

IgG, IgM, IgD, IgE, IgY, and IgA antibodies of the present invention are also

30   usefully obtained from other species, including mammals such as rodents (typically mouse, but also rat, guinea pig, and hamster), lagomorphs (typically rabbits), and also larger mammals, such as sheep, goats, cows, and horses; or egg laying birds or reptiles such as chickens or alligators. In such cases, as with the transgenic human-antibody-

producing non-human mammals, fortuitous immunization is not required, and the non-human mammal is typically affirmatively immunized, according to standard immunization protocols, with the polypeptide of the present invention. One form of avian antibodies may be generated using techniques described in WO 00/29444, published 25 May 2000.

5 As discussed above, virtually all fragments of 8 or more contiguous amino acids of a polypeptide of the present invention can be used effectively as immunogens when conjugated to a carrier, typically a protein such as bovine thyroglobulin, keyhole limpet hemocyanin, or bovine serum albumin, conveniently using a bifunctional linker such as those described elsewhere above, which discussion is incorporated by reference here.

10 Immunogenicity can also be conferred by fusion of the polypeptide of the present invention to other moieties. For example, polypeptides of the present invention can be produced by solid phase synthesis on a branched polylysine core matrix; these multiple antigenic peptides (MAPs) provide high purity, increased avidity, accurate chemical definition and improved safety in vaccine development. Tam *et al.*, *Proc. Natl. Acad. Sci.* 15 *USA* 85: 5409-5413 (1988); Posnett *et al.*, *J. Biol. Chem.* 263: 1719-1725 (1988).

Protocols for immunizing non-human mammals or avian species are well-established in the art. *See* Harlow *et al.* (eds.), Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory (1998); Coligan *et al.* (eds.), Current Protocols in Immunology, John Wiley & Sons, Inc. (2001); Zola, Monoclonal Antibodies: Preparation 20 and Use of Monoclonal Antibodies and Engineered Antibody Derivatives (Basics: From Background to Bench), Springer Verlag (2000); Gross M, Speck *J.Dtsch. Tierarztl. Wochenschr.* 103: 417-422 (1996). Immunization protocols often include multiple immunizations, either with or without adjuvants such as Freund's complete adjuvant and Freund's incomplete adjuvant, and may include naked DNA immunization (Moss, *Semin.* 25 *Immunol.* 2: 317-327 (1990).

Antibodies from non-human mammals and avian species can be polyclonal or monoclonal, with polyclonal antibodies having certain advantages in immunohistochemical detection of the polypeptides of the present invention and monoclonal antibodies having advantages in identifying and distinguishing particular 30 epitopes of the polypeptides of the present invention. Antibodies from avian species may have particular advantage in detection of the polypeptides of the present invention, in human serum or tissues (Vikinge et al., *Biosens. Bioelectron.* 13: 1257-1262 (1998). Following immunization, the antibodies of the present invention can be obtained using any

art-accepted technique. Such techniques are well known in the art and are described in detail in references such as Coligan, *supra*; Zola, *supra*; Howard *et al.* (eds.), <u>Basic Methods in Antibody Production and Characterization</u>, CRC Press (2000); Harlow, *supra*; Davis (ed.), <u>Monoclonal Antibody Protocols</u>, Vol. 45, Humana Press (1995); Delves (ed.),

5   <u>Antibody Production: Essential Techniques</u>, John Wiley & Son Ltd (1997); and Kenney, <u>Antibody Solution: An Antibody Methods Manual</u>, Chapman & Hall (1997).

Briefly, such techniques include, *inter alia*, production of monoclonal antibodies by hybridomas and expression of antibodies or fragments or derivatives thereof from host cells engineered to express immunoglobulin genes or fragments thereof. These two

10·   methods of production are not mutually exclusive: genes encoding antibodies specific for the polypeptides of the present invention can be cloned from hybridomas and thereafter expressed in other host cells. Nor need the two necessarily be performed together: *e.g.*, genes encoding antibodies specific for the polypeptides of the present invention can be cloned directly from B cells known to be specific for the desired protein, as further

15   described in U.S. Patent No. 5,627,052, the disclosure of which is incorporated herein by reference in its entirety, or from antibody-displaying phage.

Recombinant expression in host cells is particularly useful when fragments or derivatives of the antibodies of the present invention are desired.

Host cells for recombinant antibody production of whole antibodies, antibody

20   fragments, or antibody derivatives can be prokaryotic or eukaryotic.

Prokaryotic hosts are particularly useful for producing phage displayed antibodies of the present invention.

The technology of phage-displayed antibodies, in which antibody variable region fragments are fused, for example, to the gene III protein (pIII) or gene VIII protein (pVIII)·

25   for display on the surface of filamentous phage, such as M13, is by now well-established. *See, e.g.*, Sidhu, *Curr. Opin. Biotechnol.* 11(6): 610-6 (2000); Griffiths *et al.*, *Curr. Opin. Biotechnol.* 9(1): 102-8 (1998); Hoogenboom *et al.*, *Immunotechnology*, 4(1): 1-20 (1998); Rader *et al.*, *Current Opinion in Biotechnology* 8: 503-508 (1997); Aujame *et al.*, *Human Antibodies* 8: 155-168 (1997); Hoogenboom, *Trends in Biotechnol.* 15: 62-70 (1997); de

30   Kruif *et al.*, 17: 453-455 (1996); Barbas *et al.*, *Trends in Biotechnol.* 14: 230-234 (1996); Winter *et al.*, *Ann. Rev. Immunol.* 433-455 (1994). Techniques and protocols required to generate, propagate, screen (pan), and use the antibody fragments from such libraries have recently been compiled. *See, e.g.*, Barbas (2001), *supra*; Kay, *supra*; and Abelson, *supra*.

Typically, phage-displayed antibody fragments are scFv fragments or Fab fragments; when desired, full length antibodies can be produced by cloning the variable regions from the displaying phage into a complete antibody and expressing the full length antibody in a further prokaryotic or a eukaryotic host cell. Eukaryotic cells are also useful

5   for expression of the antibodies, antibody fragments, and antibody derivatives of the present invention. For example, antibody fragments of the present invention can be produced in *Pichia pastoris* and in *Saccharomyces cerevisiae*. *See, e.g.,* Takahashi *et al., Biosci. Biotechnol. Biochem.* 64(10): 2138-44 (2000); Freyre *et al.,* J. Biotechnol. 76(2-3):1 57-63 (2000); Fischer *et al., Biotechnol. Appl. Biochem.* 30 (Pt 2): 117-20

10   (1999); Pennell *et al., Res. Immunol.* 149(6): 599-603 (1998); Eldin *et al., J. Immunol. Methods.* 201(1): 67-75 (1997);, Frenken *et al., Res. Immunol.* 149(6): 589-99 (1998); and Shusta *et al., Nature Biotechnol.* 16(8): 773-7 (1998).

    · Antibodies, including antibody fragments and derivatives, of the present invention can also be produced in insect cells. *See, e.g.,* Li *et al., Protein Expr. Purif.* 21(1): 121-8

15   (2001); Ailor *et al., Biotechnol. Bioeng.* 58(2-3): 196-203 (1998); Hsu *et al., Biotechnol. Prog.* 13(1): 96-104 (1997); Edelman *et al., Immunology* 91(1): 13-9 (1997); and Nesbit *et al., J. Immunol. Methods* 151(1-2): 201-8 (1992).

Antibodies and fragments and derivatives thereof of the present invention can also be produced in plant cells, particularly maize or tobacco, Giddings *et al., Nature*

20   *Biotechnol.* 18(11): 1151-5 (2000); Gavilondo *et al., Biotechniques* 29(1): 128-38 (2000); Fischer *et al., J. Biol. Regul. Homeost. Agents* 14(2): 83-92 (2000); Fischer *et al., Biotechnol. Appl. Biochem.* 30 (Pt 2): 113-6 (1999); Fischer *et al., Biol. Chem.* 380(7-8): 825-39 (1999); Russell, *Curr. Top. Microbiol. Immunol.* 240: 119-38 (1999); and Ma *et al., Plant Physiol.* 109(2): 341-6 (1995).

25   Antibodies, including antibody fragments and derivatives, of the present invention can also be produced in transgenic, non-human, mammalian milk. *See, e.g.* Pollock et al., *J. Immunol Methods.* 231: 147-57 (1999); Young et al., *Res. Immunol.* 149: 609-10 (1998); and Limonta et al., *Immunotechnology* 1: 107-13 (1995).

Mammalian cells useful for recombinant expression of antibodies, antibody

30   fragments, and antibody derivatives of the present invention include CHO cells, COS cells, 293 cells, and myeloma cells. Verma *et al., J. Immunol. Methods* 216(1-2):165-81 (1998) review and compare bacterial, yeast, insect and mammalian expression systems for expression of antibodies. Antibodies of the present invention can also be prepared by cell

free translation, as further described in Merk *et al., J. Biochem.* (Tokyo) 125(2): 328-33 (1999) and Ryabova *et al., Nature Biotechnol.* 15(1): 79-84 (1997), and in the milk of transgenic animals, as further described in Pollock *et al., J. Immunol. Methods* 231(1-2): 147-57 (1999).

5   The invention further provides antibody fragments that bind specifically to one or more of the polypeptides of the present invention, to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited by one or more of the polypeptides of the present invention or one or more of the polypeptides encoded by the isolated nucleic acid

10   molecules of the present invention. Among such useful fragments are Fab, Fab', Fv, F(ab)'$_2$, and single chain Fv (scFv) fragments. Other useful fragments are described in Hudson, *Curr. Opin. Biotechnol.* 9(4): 395-402 (1998).

The present invention also relates to antibody derivatives that bind specifically to one or more of the polypeptides of the present invention, to one or more of the

15   polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited by one or more of the polypeptides of the present invention or one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention.

Among such useful derivatives are chimeric, primatized, and humanized

20   antibodies; such derivatives are less immunogenic in human beings, and thus are more suitable for *in vivo* administration, than are unmodified antibodies from non-human mammalian species. Another useful method is PEGylation to increase the serum half life of the antibodies.

Chimeric antibodies typically include heavy and/or light chain variable regions

25   (including both CDR and framework residues) of immunoglobulins of one species, typically mouse, fused to constant regions of another species, typically human. *See, e.g.,* Morrison *et al., Proc. Natl. Acad. Sci USA.*81(21): 6851-5 (1984); Sharon *et al., Nature* 309(5966): 364-7 (1984); Takeda *et al., Nature* 314(6010): 452-4 (1985); and U.S. Patent No. 5,807,715 the disclosure of which is incorporated herein by reference in its entirety.

30   Primatized and humanized antibodies typically include heavy and/or light chain CDRs from a murine antibody grafted into a non-human primate or human antibody V region framework, usually further comprising a human constant region, Riechmann *et al., Nature* 332(6162): 323-7 (1988); Co *et al., Nature* 351(6326): 501-2 (1991); and U.S. Patent Nos.

6,054,297; 5,821,337; 5,770,196; 5,766,886; 5,821,123; 5,869,619; 6,180,377; 6,013,256; 5,693,761; and 6,180,370, the disclosures of which are incorporated herein by reference in their entireties. Other useful antibody derivatives of the invention include heteromeric antibody complexes and antibody fusions, such as diabodies (bispecific antibodies),

5   single-chain diabodies, and intrabodies.

It is contemplated that the nucleic acids encoding the antibodies of the present invention can be operably joined to other nucleic acids forming a recombinant vector for cloning or for expression of the antibodies of the invention. Accordingly, the present invention includes any recombinant vector containing the coding sequences, or part

10   thereof, whether for eukaryotic transduction, transfection or gene therapy. Such vectors may be prepared using conventional molecular biology techniques, known to those with skill in the art, and would comprise DNA encoding sequences for the immunoglobulin V-regions including framework and CDRs or parts thereof, and a suitable promoter either with or without a signal sequence for intracellular transport. Such vectors may be

15   transduced or transfected into eukaryotic cells or used for gene therapy (Marasco et al., _Proc. Natl. Acad. Sci. (USA)_ 90: 7889-7893 (1993); Duan et al., _Proc. Natl. Acad. Sci._ _(USA)_ 91: 5075-5079 (1994), by conventional techniques, known to those with skill in the art.

The antibodies of the present invention, including fragments and derivatives

20   thereof, can usefully be labeled. It is, therefore, another aspect of the present invention to provide labeled antibodies that bind specifically to one or more of the polypeptides of the present invention, to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited by one or more of the polypeptides of the present invention or one or more of the

25   polypeptides encoded by the isolated nucleic acid molecules of the present invention. The choice of label depends, in part, upon the desired use.

For example, when the antibodies of the present invention are used for immunohistochemical staining of tissue samples, the label can usefully be an enzyme that catalyzes production and local deposition of a detectable product. Enzymes typically

30   conjugated to antibodies to permit their immunohistochemical visualization are well known, and include alkaline phosphatase, β-galactosidase, glucose oxidase, horseradish peroxidase (HRP), and urease. Typical substrates for production and deposition of visually detectable products include o-nitrophenyl-beta-D-galactopyranoside (ONPG);

o-phenylenediamine dihydrochloride (OPD); p-nitrophenyl phosphate (PNPP); p-nitrophenyl-beta-D-galactopryanoside (PNPG); 3',3'-diaminobenzidine (DAB); 3-amino-9-ethylcarbazole (AEC); 4-chloro-1-naphthol (CN); 5-bromo-4-chloro-3-indolyl-phosphate (BCIP); ABTS®; BluoGal; iodonitrotetrazolium

5   (INT); nitroblue tetrazolium chloride (NBT); phenazine methosulfate (PMS); phenolphthalein monophosphate (PMP); tetramethyl benzidine (TMB); tetranitroblue tetrazolium (TNBT); X-Gal; X-Gluc; and X-Glucoside.

Other substrates can be used to produce products for local deposition that are luminescent. For example, in the presence of hydrogen peroxide (H_2O_2), horseradish

10   peroxidase (HRP) can catalyze the oxidation of cyclic diacylhydrazides, such as luminol. Immediately following the oxidation, the luminol is in an excited state (intermediate reaction product), which decays to the ground state by emitting light. Strong enhancement of the light emission is produced by enhancers, such as phenolic compounds. Advantages include high sensitivity, high resolution, and rapid detection without radioactivity and

15   requiring only small amounts of antibody. *See, e.g.*, Thorpe *et al.*, *Methods Enzymol.* 133: 331-53 (1986); Kricka *et al.*, *J. Immunoassay* 17(1): 67-83 (1996); and Lundqvist *et al.*, *J. Biolumin. Chemilumin.* 10(6): 353-9 (1995). Kits for such enhanced chemiluminescent detection (ECL) are available commercially. The antibodies can also be labeled using colloidal gold.

20   As another example, when the antibodies of the present invention are used, *e.g.*, for flow cytometric detection, for scanning laser cytometric detection, or for fluorescent immunoassay, they can usefully be labeled with fluorophores. There are a wide variety of fluorophore labels that can usefully be attached to the antibodies of the present invention. For flow cytometric applications, both for extracellular detection and for intracellular

25   detection, common useful fluorophores can be fluorescein isothiocyanate (FITC), allophycocyanin (APC), R-phycoerythrin (PE), peridinin chlorophyll protein (PerCP), Texas Red, Cy3, Cy5, fluorescence resonance energy tandem fluorophores such as PerCP-Cy5.5, PE-Cy5, PE-Cy5.5, PE-Cy7, PE-Texas Red, and APC-Cy7.

Other fluorophores include, *inter alia*, Alexa Fluor® 350, Alexa Fluor® 488,

30   Alexa Fluor® 532, Alexa Fluor® 546, Alexa Fluor® 568, Alexa Fluor® 594, Alexa Fluor® 647 (monoclonal antibody labeling kits available from Molecular Probes, Inc., Eugene, OR, USA), BODIPY dyes, such as BODIPY 493/503, BODIPY FL, BODIPY R6G, BODIPY 530/550, BODIPY TMR, BODIPY 558/568, BODIPY 558/568, BODIPY

564/570, BODIPY 576/589, BODIPY 581/591, BODIPY TR, BODIPY 630/650,

BODIPY 650/665, Cascade Blue, Cascade Yellow, Dansyl, lissamine rhodamine B,

Marina Blue, Oregon Green 488, Oregon Green 514, Pacific Blue, rhodamine 6G,

rhodamine green, rhodamine red, tetramethylrhodamine, Texas Red (available from

5   Molecular Probes, Inc., Eugene, OR, USA), and Cy2, Cy3, Cy3.5, Cy5, Cy5.5, Cy7, all of

which are also useful for fluorescently labeling the antibodies of the present invention.

For secondary detection using labeled avidin, streptavidin, captavidin or neutravidin, the

antibodies of the present invention can usefully be labeled with biotin.

When the antibodies of the present invention are used, $e.g.$, for western blotting

10   applications, they can usefully be labeled with radioisotopes, such as ^{33}P, ^{32}P, ^{35}S, ^{3}H, and

^{125}I. As another example, when the antibodies of the present invention are used for

radioimmunotherapy, the label can usefully be ^{228}Th, ^{227}Ac, ^{225}Ac, ^{223}Ra, ^{213}Bi, ^{212}Pb,

^{212}Bi, ^{211}At, ^{203}Pb, ^{194}Os, ^{188}Re, ^{186}Re, ^{153}Sm, ^{149}Tb, ^{131}I, ^{125}I, ^{111}In, ^{105}Rh, ^{99m}Tc, ^{97}Ru, ^{90}Y,

^{90}Sr, ^{88}Y, ^{72}Se, ^{67}Cu, or ^{47}Sc.

15   As another example, when the antibodies of the present invention are to be used

for in $vivo$ diagnostic use, they can be rendered detectable by conjugation to MRI contrast

agents, such as gadolinium diethylenetriaminepentaacetic acid (DTPA), Lauffer et $al.$,

$Radiology$ 207(2): 529-38 (1998), or by radioisotopic labeling.

As would be understood, use of the labels described above is not restricted to the

20   application as for which they were mentioned.

The antibodies of the present invention, including fragments and derivatives

thereof, can also be conjugated to toxins, in order to target the toxin's ablative action to

cells that display and/or express the polypeptides of the present invention. Commonly, the

antibody in such immunotoxins is conjugated to Pseudomonas exotoxin A, diphtheria

25   toxin, shiga toxin A, anthrax toxin lethal factor, or ricin. See Hall (ed.), Immunotoxin

Methods and Protocols (Methods in Molecular Biology, vol. 166), Humana Press (2000);

and Frankel et $al.$ (eds.), Clinical Applications of Immunotoxins, Springer-Verlag (1998).

The antibodies of the present invention can usefully be attached to a substrate, and

it is, therefore, another aspect of the invention to provide antibodies that bind specifically

30   to one or more of the polypeptides of the present invention, to one or more of the

polypeptides encoded by the isolated nucleic acid molecules of the present invention, or

the binding of which can be competitively inhibited by one or more of the polypeptides of

the present invention or one or more of the polypeptides encoded by the isolated nucleic

acid molecules of the present invention, attached to a substrate. Substrates can be porous or nonporous, planar or nonplanar. For example, the antibodies of the present invention can usefully be conjugated to filtration media, such as NHS-activated Sepharose or CNBr-activated Sepharose for purposes of immunoaffinity chromatography. For example, the

5 antibodies of the present invention can usefully be attached to paramagnetic microspheres, typically by biotin-streptavidin interaction, which microsphere can then be used for isolation of cells that express or display the polypeptides of the present invention. As another example, the antibodies of the present invention can usefully be attached to the surface of a microtiter plate for ELISA.

10 As noted above, the antibodies of the present invention can be produced in prokaryotic and eukaryotic cells. It is, therefore, another aspect of the present invention to provide cells that express the antibodies of the present invention, including hybridoma cells, B cells, plasma cells, and host cells recombinantly modified to express the antibodies of the present invention.

15 In yet a further aspect, the present invention provides aptamers evolved to bind specifically to one or more of the OSPs of the present invention or to polypeptides encoded by the OSNAs of the invention.

In sum, one of skill in the art, provided with the teachings of this invention, has available a variety of methods which may be used to alter the biological properties of the

20 antibodies of this invention including methods which would increase or decrease the stability or half-life, immunogenicity, toxicity, affinity or yield of a given antibody molecule, or to alter it in any other way that may render it more suitable for a particular application.

Transgenic Animals and Cells

25 In another aspect, the invention provides transgenic cells and non-human organisms comprising nucleic acid molecules of the invention. In a preferred embodiment, the transgenic cells and non-human organisms comprise a nucleic acid molecule encoding a OSP. In a preferred embodiment, the OSP comprises an amino acid sequence selected from SEQ ID NO: 100-192, or a fragment, mutein, homologous protein

30 or allelic variant thereof. In another preferred embodiment, the transgenic cells and non-human organism comprise a OSNA of the invention, preferably a OSNA comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-99, or a part,

substantially similar nucleic acid molecule, allelic variant or hybridizing nucleic acid molecule thereof.

In another embodiment, the transgenic cells and non-human organisms have a targeted disruption or replacement of the endogenous orthologue of the human OSG. The transgenic cells can be embryonic stem cells or somatic cells. The transgenic non-human organisms can be chimeric, nonchimeric heterozygotes, and nonchimeric homozygotes. Methods of producing transgenic animals are well known in the art. *See, e.g.,* Hogan *et al.,* Manipulating the Mouse Embryo: A Laboratory Manual, 2d ed., Cold Spring Harbor Press (1999); Jackson *et al.,* Mouse Genetics and Transgenics: A Practical Approach, Oxford University Press (2000); and Pinkert, Transgenic Animal Technology: A Laboratory Handbook, Academic Press (1999).

Any technique known in the art may be used to introduce a nucleic acid molecule of the invention into an animal to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection. (*see, e.g.,* Paterson *et al., Appl. Microbiol. Biotechnol.* 40: 691-698 (1994); Carver *et al., Biotechnology* 11: 1263-1270 (1993); Wright *et al., Biotechnology* 9: 830-834 (1991); and U.S. Patent No. 4,873,191, herein incorporated by reference in its entirety); retrovirus-mediated gene transfer into germ lines, blastocysts or embryos (*see, e.g.,* Van der Putten *et al., Proc. Natl. Acad. Sci., USA* 82: 6148-6152 (1985)); gene targeting in embryonic stem cells (*see, e.g.,* Thompson *et al., Cell* 56: 313-321 (1989)); electroporation of cells or embryos (*see, e.g.,* Lo, 1983, *Mol. Cell. Biol.* 3: 1803-1814 (1983)); introduction using a gene gun (*see, e.g.,* Ulmer *et al., Science* 259: 1745-49 (1993); introducing nucleic acid constructs into embryonic pleuripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (*see, e.g.,* Lavitrano *et al., Cell* 57: 717-723 (1989)).

Other techniques include, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (*see, e.g.,* Campell *et al., Nature* 380: 64-66 (1996); Wilmut *et al., Nature* 385: 810-813 (1997)). The present invention provides for transgenic animals that carry the transgene (*i.e.,* a nucleic acid molecule of the invention) in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e. e., mosaic animals or chimeric animals.

The transgene may be integrated as a single transgene or as multiple copies, such as in concatamers, e. g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following,

e.g., the teaching of Lasko *et al.* *et al.*, *Proc. Natl. Acad. Sci. USA* 89: 6232- 6236 (1992).
The regulatory sequences required for such a cell-type specific activation will depend
upon the particular cell type of interest, and will be apparent to those of skill in the art.

5   Once transgenic animals have been generated, the expression of the recombinant
gene may be assayed utilizing standard techniques. Initial screening may be accomplished
by Southern blot analysis or PCR techniques to analyze animal tissues to verify that
integration of the transgene has taken place. The level of mRNA expression of the
transgene in the tissues of the transgenic animals may also be assessed using techniques
which include, but are not limited to, Northern blot analysis of tissue samples obtained
10   from the animal, in situ hybridization analysis, and reverse transcriptase-PCR (RT-PCR).
Samples of transgenic gene-expressing tissue may also be evaluated
immunocytochemically or immunohistochemically using antibodies specific for the
transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or
15   crossbred to produce colonies of the particular animal. Examples of such breeding
strategies include, but are not limited to: outbreeding of founder animals with more than
one integration site in order to establish separate lines; inbreeding of separate lines in
order to produce compound transgenics that express the transgene at higher levels because
of the effects of additive expression of each transgene; crossing of heterozygous
20   transgenic animals to produce animals homozygous for a given integration site in order to
both augment expression and eliminate the need for screening of animals by DNA
analysis; crossing of separate homozygous lines to produce compound heterozygous or
homozygous lines; and breeding to place the transgene on a distinct background that is
appropriate for an experimental model of interest.

25   Transgenic animals of the invention have uses which include, but are not limited
to, animal model systems useful in elaborating the biological function of polypeptides of
the present invention, studying conditions and/or disorders associated with aberrant
expression, and in screening for compounds effective in ameliorating such conditions
and/or disorders.

30   Methods for creating a transgenic animal with a disruption of a targeted gene are
also well known in the art. In general, a vector is designed to comprise some nucleotide
sequences homologous to the endogenous targeted gene. The vector is introduced into a
cell so that it may integrate, via homologous recombination with chromosomal sequences,

into the endogenous gene, thereby disrupting the function of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type. See, e.g., Gu et al., Science 265: 103-106 (1994). The regulatory sequences required for such a cell-type specific inactivation will

5   depend upon the particular cell type of interest, and will be apparent to those of skill in the art. See, e.g., Smithies et al., Nature 317: 230-234 (1985); Thomas et al., Cell 51: 503-512 (1987); Thompson et al., Cell 5: 313-321 (1989).

In one embodiment, a mutant, non-functional nucleic acid molecule of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the

10   endogenous nucleic acid sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention in vivo. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via

15   targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene. See, e.g., Thomas, supra and Thompson, supra. However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly

20   administered or targeted to the required site in vivo using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e.g., knockouts) are administered to a

25   patient in vivo. Such cells may be obtained from an animal or patient or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e.g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered in vitro using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt

30   the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures,

including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, *e.g.*, in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, *e.g.*, genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. *See, e.g.*, U.S. Patent Nos. 5,399,349 and 5,460,959, each of which is incorporated by reference herein in its entirety.

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Computer Readable Means

A further aspect of the invention is a computer readable means for storing the nucleic acid and amino acid sequences of the instant invention. In a preferred embodiment, the invention provides a computer readable means for storing SEQ ID NO: 100-192 and SEQ ID NO: 1-99 as described herein, as the complete set of sequences or in any combination. The records of the computer readable means can be accessed for reading and display and for interface with a computer system for the application of programs allowing for the location of data upon a query for data meeting certain criteria,

the comparison of sequences, the alignment or ordering of sequences meeting a set of

criteria, and the like.

The nucleic acid and amino acid sequences of the invention are particularly useful

as components in databases useful for search analyses as well as in sequence analysis

5   algorithms. As used herein, the terms "nucleic acid sequences of the invention" and

"amino acid sequences of the invention" mean any detectable chemical or physical

characteristic of a polynucleotide or polypeptide of the invention that is or may be reduced

to or stored in a computer readable form. These include, without limitation,

chromatographic scan data or peak data, photographic data or scan data therefrom, and

10   mass spectrographic data.

This invention provides computer readable media having stored thereon sequences

of the invention. A computer readable medium may comprise one or more of the

following: a nucleic acid sequence comprising a sequence of a nucleic acid sequence of

the invention; an amino acid sequence comprising an amino acid sequence of the

15   invention; a set of nucleic acid sequences wherein at least one of said sequences comprises

the sequence of a nucleic acid sequence of the invention; a set of amino acid sequences

wherein at least one of said sequences comprises the sequence of an amino acid sequence

of the invention; a data set representing a nucleic acid sequence comprising the sequence

of one or more nucleic acid sequences of the invention; a data set representing a nucleic

20   acid sequence encoding an amino acid sequence comprising the sequence of an amino acid

sequence of the invention; a set of nucleic acid sequences wherein at least one of said

sequences comprises the sequence of a nucleic acid sequence of the invention; a set of

amino acid sequences wherein at least one of said sequences comprises the sequence of an

amino acid sequence of the invention; a data set representing a nucleic acid sequence

25   comprising the sequence of a nucleic acid sequence of the invention; a data set

representing a nucleic acid sequence encoding an amino acid sequence comprising the

sequence of an amino acid sequence of the invention. The computer readable medium can

be any composition of matter used to store information or data, including, for example,

commercially available floppy disks, tapes, hard drives, compact disks, and video disks.

30   Also provided by the invention are methods for the analysis of character

sequences, particularly genetic sequences. Preferred methods of sequence analysis

include, for example, methods of sequence homology analysis, such as identity and

similarity analysis, RNA structure analysis, sequence assembly, cladistic analysis,

sequence motif analysis, open reading frame determination, nucleic acid base calling, and sequencing chromatogram peak analysis.

A computer-based method is provided for performing nucleic acid sequence identity or similarity identification. This method comprises the steps of providing a

5   nucleic acid sequence comprising the sequence of a nucleic acid of the invention in a computer readable medium; and comparing said nucleic acid sequence to at least one nucleic acid or amino acid sequence to identify sequence identity or similarity.

A computer-based method is also provided for performing amino acid homology identification, said method comprising the steps of: providing an amino acid sequence

10   comprising the sequence of an amino acid of the invention in a computer readable medium; and comparing said amino acid sequence to at least one nucleic acid or an amino acid sequence to identify homology.

A computer-based method is still further provided for assembly of overlapping nucleic acid sequences into a single nucleic acid sequence, said method comprising the

15   steps of: providing a first nucleic acid sequence comprising the sequence of a nucleic acid of the invention in a computer readable medium; and screening for at least one overlapping region between said first nucleic acid sequence and a second nucleic acid sequence. In addition, the invention includes a·method of using patterns of expression associated with either the nucleic acids or proteins in a computer-based method to

20   diagnose disease.

Diagnostic Methods for ovarian Cancer

The present invention also relates to quantitative and qualitative diagnostic assays and methods for detecting, diagnosing, monitoring, staging and predicting cancers by comparing expression of a OSNA or a OSP in a human patient that has or may have

25   ovarian cancer, or who is at risk of developing ovarian cancer, with the expression of a OSNA or a OSP in a normal human control. For purposes of the present invention, "expression of a OSNA" or "OSNA expression" means the quantity of OSNA mRNA that can be measured by any method known in the art or the level of transcription that can be measured by any method known in the art in a cell, tissue, organ or whole patient.

30   Similarly, the term "expression of a OSP" or "OSP expression" means the amount of OSP that can be measured by any method known in the art or the level of translation of a OSNA that can be measured by any method known in the art.

The present invention provides methods for diagnosing ovarian cancer in a patient, by analyzing for changes in levels of OSNA or OSP in cells, tissues, organs or bodily fluids compared with levels of OSNA or OSP in cells, tissues, organs or bodily fluids of preferably the same type from a normal human control, wherein an increase, or decrease in

5   certain cases, in levels of a OSNA or OSP in the patient versus the normal human control is associated with the presence of ovarian cancer or with a predilection to the disease. In another preferred embodiment, the present invention provides methods for diagnosing ovarian cancer in a patient by analyzing changes in the structure of the mRNA of a OSG compared to the mRNA from a normal control. These changes include, without limitation, .

10   aberrant splicing, alterations in polyadenylation and/or alterations in 5' nucleotide capping. In yet another preferred embodiment, the present invention provides methods for · diagnosing ovarian cancer in a patient by analyzing changes in a OSP compared to a OSP from a normal patient. These changes include, *e.g.*, alterations, including post translational modifications such as glycosylation and/or phosphorylation of the OSP or

15   changes in the subcellular OSP localization.

For purposes of the present invention, diagnosing means that OSNA or OSP levels are used to determine the presence or absence of disease in a patient. As will be understood by those of skill in the art, measurement of other diagnostic parameters may be · required for definitive diagnosis or determination of the appropriate treatment for the

20   disease. The determination may be made by a clinician, a doctor, a testing laboratory, or a patient using an over the counter test. The patient may have symptoms of disease or may be asymptomatic. In addition, the OSNA or OSP levels of the present invention may be used as screening marker to determine whether further tests or biopsies are warranted. In addition, the OSNA or OSP levels may be used to determine the vulnerability or

25   susceptibility to disease.

In a preferred embodiment, the expression of a OSNA is measured by determining the amount of a mRNA that encodes an amino acid sequence selected from SEQ ID NO: 100-192, a homolog, an allelic variant, or a fragment thereof. In a more preferred embodiment, the OSNA expression that is measured is the level of expression of a OSNA

30   mRNA selected from SEQ ID NO: 1-99, or a hybridizing nucleic acid, homologous nucleic acid or allelic variant thereof, or a part of any of these nucleic acid molecules. OSNA expression may be measured by any method known in the art, such as those described *supra*, including measuring mRNA expression by Northern blot, quantitative or

qualitative reverse transcriptase PCR (RT-PCR), microarray, dot or slot blots or *in situ* hybridization. *See, e.g.,* Ausubel (1992), *supra*; Ausubel (1999), *supra*; Sambrook (1989), *supra*; and Sambrook (2001), *supra*. OSNA transcription may be measured by any method known in the art including using a reporter gene hooked up to the promoter of

5   a OSG of interest or doing nuclear run-off assays. Alterations in mRNA structure, *e.g.*, aberrant splicing variants, may be determined by any method known in the art, including, RT-PCR followed by sequencing or restriction analysis. As necessary, OSNA expression may be compared to a known control, such as normal ovarian nucleic acid, to detect a change in expression.

10   In another preferred embodiment, the expression of a OSP is measured by determining the level of a OSP having an amino acid sequence selected from the group consisting of SEQ ID NO: 100-192, a homolog, an allelic variant, or a fragment thereof. Such levels are preferably determined in at least one of cells, tissues, organs and/or bodily fluids, including determination of normal and abnormal levels. Thus, for instance, a

15   diagnostic assay in accordance with the invention for diagnosing over- or underexpression of a OSNA or OSP compared to normal control bodily fluids, cells, or tissue samples may be used to diagnose the presence of ovarian cancer. The expression level of a OSP may be determined by any method known in the art, such as those described *supra*. In a preferred embodiment, the OSP expression level may be determined by radioimmunoassays,

20   competitive-binding assays, ELISA, Western blot, FACS, immunohistochemistry, immunoprecipitation, proteomic approaches: two-dimensional gel electrophoresis (2D electrophoresis) and non-gel-based approaches such as mass spectrometry or protein interaction profiling. *See, e.g,* Harlow (1999), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), *supra*. Alterations in the OSP structure may be determined by any method known

25   in the art, including, *e.g.*, using antibodies that specifically recognize phosphoserine, phosphothreonine or phosphotyrosine residues, two-dimensional polyacrylamide gel electrophoresis (2D PAGE) and/or chemical analysis of amino acid residues of the protein. *Id.*

          In a preferred embodiment, a radioimmunoassay (RIA) or an ELISA is used. An

30   antibody specific to a OSP is prepared if one is not already available. In a preferred embodiment, the antibody is a monoclonal antibody. The anti-OSP antibody is bound to a solid support and any free protein binding sites on the solid support are blocked with a protein such as bovine serum albumin. A sample of interest is incubated with the antibody

on the solid support under conditions in which the OSP will bind to the anti-OSP antibody. The sample is removed, the solid support is washed to remove unbound material, and an anti-OSP antibody that is linked to a detectable reagent (a radioactive substance for RIA and an enzyme for ELISA) is added to the solid support and incubated under conditions in

5 which binding of the OSP to the labeled antibody will occur. After binding, the unbound labeled antibody is removed by washing. For an ELISA, one or more substrates are added to produce a colored reaction product that is based upon the amount of an OSP in the sample. For an RIA, the solid support is counted for radioactive decay signals by any method known in the art. Quantitative results for both RIA and ELISA typically are

10 obtained by reference to a standard curve.

Other methods to measure OSP levels are known in the art. For instance, a competition assay may be employed wherein an anti-OSP antibody is attached to a solid support and an allocated amount of a labeled OSP and a sample of interest are incubated with the solid support. The amount of labeled OSP attached to the solid support can be

15 correlated to the quantity of a OSP in the sample.

Of the proteomic approaches, 2D PAGE is a well known technique. Isolation of individual proteins from a sample such as serum is accomplished using sequential separation of proteins by isoelectric point and molecular weight. Typically, polypeptides are first separated by isoelectric point (the first dimension) and then separated by size

20 using an electric current (the second dimension). In general, the second dimension is perpendicular to the first dimension. Because no two proteins with different sequences are identical on the basis of both size and charge, the result of 2D PAGE is a roughly square gel in which each protein occupies a unique spot. Analysis of the spots with chemical or antibody probes, or subsequent protein microsequencing can reveal the relative abundance

25 of a given protein and the identity of the proteins in the sample.

Expression levels of a OSNA can be determined by any method known in the art, including PCR and other nucleic acid methods, such as ligase chain reaction (LCR) and nucleic acid sequence based amplification (NASBA), can be used to detect malignant cells for diagnosis and monitoring of various malignancies. For example, reverse-transcriptase

30 PCR (RT-PCR) is a powerful technique which can be used to detect the presence of a specific mRNA population in a complex mixture of thousands of other mRNA species. In RT-PCR, an mRNA species is first reverse transcribed to complementary DNA (cDNA)

with use of the enzyme reverse transcriptase; the cDNA is then amplified as in a standard PCR reaction.

Hybridization to specific DNA molecules (*e.g.*, oligonucleotides) arrayed on a solid support can be used to both detect the expression of and quantitate the level of

5 expression of one or more OSNAs of interest. In this approach, all or a portion of one or more OSNAs is fixed to a substrate. A sample of interest, which may comprise RNA, *e.g.*, total RNA or polyA-selected mRNA, or a complementary DNA (cDNA) copy of the RNA is incubated with the solid support under conditions in which hybridization will occur between the DNA on the solid support and the nucleic acid molecules in the sample of

10 interest. Hybridization between the substrate-bound DNA and the nucleic acid molecules in the sample can be detected and quantitated by several means, including, without limitation, radioactive labeling or fluorescent labeling of the nucleic acid molecule or a secondary molecule designed to detect the hybrid.

The above tests can be carried out on samples derived from a variety of cells,

15 bodily fluids and/or tissue extracts such as homogenates or solubilized tissue obtained from a patient. Tissue extracts are obtained routinely from tissue biopsy and autopsy material. Bodily fluids useful in the present invention include blood, urine, saliva or any other bodily secretion or derivative thereof. As used herein "blood" includes whole blood, plasma, serum, circulating epithelial cells, constituents, or any derivative of blood.

20 In addition to detection in bodily fluids, the proteins and nucleic acids of the invention are suitable to detection by cell capture technology. Whole cells may be captured by a variety methods for example magnetic separation, U.S. Patent. Nos. 5,200,084; 5,186,827; 5,108,933; 4,925,788, the disclosures of which are incorporated herein by reference in their entireties. Epithelial cells may be captured using such

25 products as Dynabeads® or CELLection™ (Dynal Biotech, Oslo, Norway). Alternatively, fractions of blood may be captured, e.g., the buffy coat fraction (50mm cells isolated from 5ml of blood) containing epithelial cells. In addition, cancer cells may be captured using the techniques described in WO 00/47998, the disclosure of which is incorporated herein by reference in its entirety. Once the cells are captured or concentrated, the proteins or

30 nucleic acids are detected by the means described in the subject application. Alternatively, nucleic acids may be captured directly from blood samples, see U.S. Patent Nos. 6,156,504, 5,501,963; or WO 01/42504 , the disclosures of which are incorporated herein by reference in their entireties.

In a preferred embodiment, the specimen tested for expression of OSNA or OSP includes without limitation ovarian tissue, ovarian cells grown in cell culture, blood, serum, lymph node tissue, and lymphatic fluid. In another preferred embodiment, especially when metastasis of a primary ovarian cancer is known or suspected, specimens

5   include, without limitation, tissues from brain, bone, bone marrow, liver, lungs, colon, and adrenal glands. In general, the tissues may be sampled by biopsy, including, without limitation, needle biopsy, e.g., transthoracic needle aspiration, cervical mediatinoscopy, endoscopic lymph node biopsy, video-assisted thoracoscopy, exploratory thoracotomy, bone marrow biopsy and bone marrow aspiration.

10   All the methods of the present invention may optionally include determining the expression levels of one or more other cancer markers in addition to determining the expression level of a OSNA or OSP. In many cases, the use of another cancer marker will decrease the likelihood of false positives or false negatives. In one embodiment, the one or more other cancer markers include other OSNA or OSPs as disclosed herein. Other

15   cancer markers useful in the present invention will depend on the cancer being tested and are known to those of skill in the art. In a preferred embodiment, at least one other cancer marker in addition to a particular OSNA or OSP is measured. In a more preferred · embodiment, at least two other additional cancer markers are used. In an even more preferred embodiment, at least three, more preferably at least five, even more preferably at ·

20   least ten additional cancer markers are used.

Diagnosing

In one aspect, the invention provides a method for determining the expression levels and/or structural alterations of one or more OSNA and/or OSP in a sample from a

25   patient suspected of having ovarian cancer. In general, the method comprises the steps of obtaining the sample from the patient, determining the expression level or structural alterations of a OSNA and/or OSP and then ascertaining whether the patient has ovarian cancer from the expression level of the OSNA or OSP. In general, if high expression relative to a control of a OSNA or OSP is indicative of ovarian cancer, a diagnostic assay

30   is considered positive if the level of expression of the OSNA or OSP is at least one and a half times higher, and more preferably are at least two times higher, still more preferably five times higher, even more preferably at least ten times higher, than in preferably the same cells, tissues or bodily fluid of a normal human control. In contrast, if low

expression relative to a control of a OSNA or OSP is indicative of ovarian cancer, a diagnostic assay is considered positive if the level of expression of the OSNA or OSP is at least one and a half times lower, and more preferably are at least two times lower, still more preferably five times lower, even more preferably at least ten times lower than in

5   preferably the same cells, tissues or bodily fluid of a normal human control. The normal human control may be from a different patient or from uninvolved tissue of the same patient.

The present invention also provides a method of determining whether ovarian cancer has metastasized in a patient. One may identify whether the ovarian cancer has

10   metastasized by measuring the expression levels and/or structural alterations of one or more OSNAs and/or OSPs in a variety of tissues. The presence of a OSNA or OSP in a certain tissue at levels higher than that of corresponding noncancerous tissue (e.g., the same tissue from another individual) is indicative of metastasis if high level expression of a OSNA or OSP is associated with ovarian cancer. Similarly, the presence of a OSNA or

15   OSP in a tissue at levels lower than that of corresponding noncancerous tissue is indicative of metastasis if low level expression of a OSNA or OSP is associated with ovarian cancer. Further, the presence of a structurally altered OSNA or OSP that is associated with ovarian cancer is also indicative of metastasis.

In general, if high expression relative to a control of a OSNA or OSP is indicative

20   of metastasis, an assay for metastasis is considered positive if the level of expression of the OSNA or OSP is at least one and a half times higher, and more preferably are at least two times higher, still more preferably five times higher, even more preferably at least ten times higher, than in preferably the same cells, tissues or bodily fluid of a normal human control. In contrast, if low expression relative to a control of a OSNA or OSP is indicative

25   of metastasis, an assay for metastasis is considered positive if the level of expression of the OSNA or OSP is at least one and a half times lower, and more preferably are at least two times lower, still more preferably five times lower, even more preferably at least ten times lower than in preferably the same cells, tissues or bodily fluid of a normal human control.

30   *Staging*

The invention also provides a method of staging ovarian cancer in a human patient. The method comprises identifying a human patient having ovarian cancer and analyzing

cells, tissues or bodily fluids from such human patient for expression levels and/or structural alterations of one or more OSNAs or OSPs. First, one or more tumors from a variety of patients are staged according to procedures well known in the art, and the expression levels of one or more OSNAs or OSPs is determined for each stage to obtain a

5   standard expression level for each OSNA and OSP. Then, the OSNA or OSP expression levels of the OSNA or OSP are determined in a biological sample from a patient whose stage of cancer is not known. The OSNA or OSP expression levels from the patient are then compared to the standard expression level. By comparing the expression level of the OSNAs and OSPs from the patient to the standard expression levels, one may determine

10   the stage of the tumor. The same procedure may be followed using structural alterations of a OSNA or OSP to determine the stage of a ovarian cancer.

Monitoring

Further provided is a method of monitoring ovarian cancer in a human patient. One may monitor a human patient to determine whether there has been metastasis and, if

15   there has been, when metastasis began to occur. One may also monitor a human patient to determine whether a preneoplastic lesion has become cancerous. One may also monitor a human patient to determine whether a therapy, *e.g.*, chemotherapy, radiotherapy or surgery, has decreased or eliminated the ovarian cancer. The monitoring may determine if there has been a reoccurrence and, if so, determine its nature. The method comprises

20   identifying a human patient that one wants to monitor for ovarian cancer, periodically analyzing cells, tissues or bodily fluids from such human patient for expression levels of one or more OSNAs or OSPs, and comparing the OSNA or OSP levels over time to those OSNA or OSP expression levels obtained previously. Patients may also be monitored by measuring one or more structural alterations in a OSNA or OSP that are associated with

25   ovarian cancer.

If increased expression of a OSNA or OSP is associated with metastasis, treatment failure, or conversion of a preneoplastic lesion to a cancerous lesion, then detecting an increase in the expression level of a OSNA or OSP indicates that the tumor is metastasizing, that treatment has failed or that the lesion is cancerous, respectively. One

30   having ordinary skill in the art would recognize that if this were the case, then a decreased expression level would be indicative of no metastasis, effective therapy or failure to progress to a neoplastic lesion. If decreased expression of a OSNA or OSP is associated

with metastasis, treatment failure, or conversion of a preneoplastic lesion to a cancerous lesion, then detecting a decrease in the expression level of a OSNA or OSP indicates that the tumor is metastasizing, that treatment has failed or that the lesion is cancerous, respectively. In a preferred embodiment, the levels of OSNAs or OSPs are determined

5   from the same cell type, tissue or bodily fluid as prior patient samples. Monitoring a patient for onset of ovarian cancer metastasis is periodic and preferably is done on a quarterly basis, but may be done more or less frequently.

The methods described herein can further be utilized as prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with

10   increased or decreased expression levels of a OSNA and/or OSP. The present invention provides a method in which a test sample is obtained from a human patient and one or more OSNAs and/or OSPs are detected. The presence of higher (or lower) OSNA or OSP levels as compared to normal human controls is diagnostic for the human patient being at risk for developing cancer, particularly ovarian cancer. The effectiveness of therapeutic

15   agents to decrease (or increase) expression or activity of one or more OSNAs and/or OSPs of the invention can also be monitored by analyzing levels of expression of the OSNAs and/or OSPs in a human patient in clinical trials or in *in vitro* screening assays such as in human cells. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the human patient or cells, as the case may be, to the agent

20   being tested.

Detection of Genetic Lesions or Mutations

The methods of the present invention can also be used to detect genetic lesions or mutations in a OSG, thereby determining if a human with the genetic lesion is susceptible to developing ovarian cancer or to determine what genetic lesions are responsible, or are

25   partly responsible, for a person's existing ovarian cancer. Genetic lesions can be detected, for example, by ascertaining the existence of a deletion, insertion and/or substitution of one or more nucleotides from the OSGs of this invention, a chromosomal rearrangement of a OSG, an aberrant modification of a OSG (such as of the methylation pattern of the genomic DNA), or allelic loss of a OSG. Methods to detect such lesions in the OSG of

30   this invention are known to those having ordinary skill in the art following the teachings of the specification.

Methods of Detecting Noncancerous ovarian Diseases

The present invention also provides methods for determining the expression levels and/or structural alterations of one or more OSNAs and/or OSPs in a sample from a patient suspected of having or known to have a noncancerous ovarian disease. In general,

5 the method comprises the steps of obtaining a sample from the patient, determining the expression level or structural alterations of a OSNA and/or OSP, comparing the expression level or structural alteration of the OSNA or OSP to a normal ovarian control, and then ascertaining whether the patient has a noncancerous ovarian disease. In general, if high expression relative to a control of a OSNA or OSP is indicative of a particular

10 noncancerous ovarian disease, a diagnostic assay is considered positive if the level of expression of the OSNA or OSP is at least two times higher, and more preferably are at least five times higher, even more preferably at least ten times higher, than in preferably the same cells, tissues or bodily fluid of a normal human control. In contrast, if low expression relative to a control of a OSNA or OSP is indicative of a noncancerous ovarian

15 disease, a diagnostic assay is considered positive if the level of expression of the OSNA or OSP is at least two times lower, more preferably are at least five times lower, even more preferably at least ten times lower than in preferably the same cells, tissues or bodily fluid of a normal human control. The normal human control may be from a different patient or from uninvolved tissue of the same patient.

20 One having ordinary skill in the art may determine whether a OSNA and/or OSP is associated with a particular noncancerous ovarian disease by obtaining ovarian tissue from a patient having a noncancerous ovarian disease of interest and determining which OSNAs and/or OSPs are expressed in the tissue at either a higher or a lower level than in normal ovarian tissue. In another embodiment, one may determine whether a OSNA or OSP

25 exhibits structural alterations in a particular noncancerous ovarian disease state by obtaining ovarian tissue from a patient having a noncancerous ovarian disease of interest and determining the structural alterations in one or more OSNAs and/or OSPs relative to normal ovarian tissue.

Methods for Identifying ovarian Tissue

30 In another aspect, the invention provides methods for identifying ovarian tissue. These methods are particularly useful in, *e.g.*, forensic science, ovarian cell differentiation and development, and in tissue engineering.

In one embodiment, the invention provides a method for determining whether a sample is ovarian tissue or has ovarian tissue-like characteristics. The method comprises the steps of providing a sample suspected of comprising ovarian tissue or having ovarian tissue-like characteristics, determining whether the sample expresses one or more OSNAs

5   and/or OSPs, and, if the sample expresses one or more OSNAs and/or OSPs, concluding that the sample comprises ovarian tissue. In a preferred embodiment, the OSNA encodes a polypeptide having an amino acid sequence selected from SEQ ID NO: 100-192, or a homolog, allelic variant or fragment thereof. In a more preferred embodiment, the OSNA has a nucleotide sequence selected from SEQ ID NO: 1-99, or a hybridizing nucleic acid,

10   an allelic variant or a part thereof. Determining whether a sample expresses a OSNA can be accomplished by any method known in the art. Preferred methods include hybridization to microarrays, Northern blot hybridization, and quantitative or qualitative RT-PCR. In another preferred embodiment, the method can be practiced by determining whether a OSP is expressed. Determining whether a sample expresses a OSP can be

15   accomplished by any method known in the art. Preferred methods include Western blot, ELISA, RIA and 2D PAGE. In one embodiment, the OSP has an amino acid sequence selected from SEQ ID NO: 100-192, or a homolog, allelic variant or fragment thereof. In another preferred embodiment, the expression of at least two OSNAs and/or OSPs is determined. In a more preferred embodiment, the expression of at least three, more

20   preferably four and even more preferably five OSNAs and/or OSPs are determined.

In one embodiment, the method can be used to determine whether an unknown tissue is ovarian tissue. This is particularly useful in forensic science, in which small, damaged pieces of tissues that are not identifiable by microscopic or other means are recovered from a crime or accident scene. In another embodiment, the method can be

25   used to determine whether a tissue is differentiating or developing into ovarian tissue. This is important in monitoring the effects of the addition of various agents to cell or tissue culture, *e.g.*, in producing new ovarian tissue by tissue engineering. These agents include, *e.g.*, growth and differentiation factors, extracellular matrix proteins and culture medium. Other factors that may be measured for effects on tissue development and

30   differentiation include gene transfer into the cells or tissues, alterations in pH, aqueous:air interface and various other culture conditions.

Methods for Producing and Modifying ovarian Tissue

In another aspect, the invention provides methods for producing engineered ovarian tissue or cells. In one embodiment, the method comprises the steps of providing cells, introducing a OSNA or a OSG into the cells, and growing the cells under conditions

5   in which they exhibit one or more properties of ovarian tissue cells. In a preferred embodiment, the cells are pleuripotent. As is well known in the art, normal ovarian tissue comprises a large number of different cell types. Thus, in one embodiment, the engineered ovarian tissue or cells comprises one of these cell types. In another embodiment, the engineered ovarian tissue or cells comprises more than one ovarian cell

10   type. Further, the culture conditions of the cells or tissue may require manipulation in order to achieve full differentiation and development of the ovarian cell tissue. Methods for manipulating culture conditions are well known in the art.

Nucleic acid molecules encoding one or more OSPs are introduced into cells, preferably pleuripotent cells. In a preferred embodiment, the nucleic acid molecules

15   encode OSPs having amino acid sequences selected from SEQ ID NO: 100-192, or homologous proteins, analogs, allelic variants or fragments thereof. In a more preferred embodiment, the nucleic acid molecules have a nucleotide sequence selected from SEQ ID NO: 1-99, or hybridizing nucleic acids, allelic variants or parts thereof. In another highly preferred embodiment, a OSG is introduced into the cells. Expression vectors and

20   methods of introducing nucleic acid molecules into cells are well known in the art and are described in detail, *supra.*

Artificial ovarian tissue may be used to treat patients who have lost some or all of their ovarian function.

Pharmaceutical Compositions

25   In another aspect, the invention provides pharmaceutical compositions comprising the nucleic acid molecules, polypeptides, fusion proteins, antibodies, antibody derivatives, antibody fragments, agonists, antagonists, or inhibitors of the present invention. In a preferred embodiment, the pharmaceutical composition comprises a OSNA or part thereof. In a more preferred embodiment, the OSNA has a nucleotide sequence selected from the

30   group consisting of SEQ ID NO: 1-99, a nucleic acid that hybridizes thereto, an allelic variant thereof, or a nucleic acid that has substantial sequence identity thereto. In another preferred embodiment, the pharmaceutical composition comprises a OSP or fragment

thereof. In a more preferred embodiment, the pharmaceutical composition comprises a OSP having an amino acid sequence that is selected from the group consisting of SEQ ID NO: 100-192, a polypeptide that is homologous thereto, a fusion protein comprising all or a portion of the polypeptide, or an analog or derivative thereof. In another preferred

5   embodiment, the pharmaceutical composition comprises an anti-OSP antibody, preferably an antibody that specifically binds to a OSP having an amino acid that is selected from the group consisting of SEQ ID NO: 100-192, or an antibody that binds to a polypeptide that is homologous thereto, a fusion protein comprising all or a portion of the polypeptide, or an analog or derivative thereof.

10   Such a composition typically contains from about 0.1 to 90% by weight of a therapeutic agent of the invention formulated in and/or with a pharmaceutically acceptable carrier or excipient.

    Pharmaceutical formulation is a well-established art that is further described in Gennaro (ed.), Remington: The Science and Practice of Pharmacy, 20th ed., Lippincott,

15   Williams & Wilkins (2000); Ansel et al., Pharmaceutical Dosage Forms and Drug Delivery Systems, 7th ed., Lippincott Williams & Wilkins (1999); and Kibbe (ed.), Handbook of Pharmaceutical Excipients American Pharmaceutical Association, 3rd ed. (2000) and thus need not be described in detail herein.

    Briefly, formulation of the pharmaceutical compositions of the present invention

20   will depend upon the route chosen for administration. The pharmaceutical compositions utilized in this invention can be administered by various routes including both enteral and parenteral routes, including oral, intravenous, intramuscular, subcutaneous, inhalation, topical, sublingual, rectal, intra-arterial, intramedullary, intrathecal, intraventricular, transmucosal, transdermal, intranasal, intraperitoneal, intrapulmonary, and intrauterine.

25   Oral dosage forms can be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

    Solid formulations of the compositions for oral administration can contain suitable carriers or excipients, such as carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other

30   plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, or microcrystalline cellulose; gums including arabic and tragacanth; proteins such as gelatin and collagen; inorganics, such as kaolin, calcium

carbonate, dicalcium phosphate, sodium chloride; and other agents such as acacia and alginic acid.

Agents that facilitate disintegration and/or solubilization can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium

5 alginate, microcrystalline cellulose, cornstarch, sodium starch glycolate, and alginic acid.

Tablet binders that can be used include acacia, methylcellulose, sodium carboxymethylcellulose, polyvinylpyrrolidone (Povidone™), hydroxypropyl methylcellulose, sucrose, starch and ethylcellulose.

Lubricants that can be used include magnesium stearates, stearic acid, silicone

10 fluid, talc, waxes, oils, and colloidal silica.

Fillers, agents that facilitate disintegration and/or solubilization, tablet binders and lubricants, including the aforementioned, can be used singly or in combination.

Solid oral dosage forms need not be uniform throughout. For example, dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar

15 solutions, which can also contain gum arabic, talc, polyvinylpyrrolidone, carbopol·gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures.

Oral dosage forms of the present invention include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or

20 sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

25 Additionally, dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Liquid formulations of the pharmaceutical compositions for oral (enteral) administration are prepared in water or other aqueous vehicles and can contain various suspending agents such as methylcellulose, alginates, tragacanth, pectin, kelgin,

30 carrageenan, acacia, polyvinylpyrrolidone, and polyvinyl alcohol. The liquid formulations can also include solutions, emulsions, syrups and elixirs containing, together with the active compound(s), wetting agents, sweeteners, and coloring and flavoring agents.

The pharmaceutical compositions of the present invention can also be formulated for parenteral administration. Formulations for parenteral administration can be in the form of aqueous or non-aqueous isotonic sterile injection solutions or suspensions.

For intravenous injection, water soluble versions of the compounds of the present invention are formulated in, or if provided as a lyophilate, mixed with, a physiologically acceptable fluid vehicle, such as 5% dextrose ("D5"), physiologically buffered saline, 0.9% saline, Hanks' solution, or Ringer's solution. Intravenous formulations may include carriers, excipients or stabilizers including, without limitation, calcium, human serum albumin, citrate, acetate, calcium chloride, carbonate, and other salts.

Intramuscular preparations, e.g. a sterile formulation of a suitable soluble salt form of the compounds of the present invention, can be dissolved and administered in a pharmaceutical excipient such as Water-for-Injection, 0.9% saline, or 5% glucose solution. Alternatively, a suitable insoluble form of the compound can be prepared and administered as a suspension in an aqueous base or a pharmaceutically acceptable oil base, such as an ester of a long chain fatty acid (e.g., ethyl oleate), fatty oils such as sesame oil, triglycerides, or liposomes.

Parenteral formulations of the compositions can contain various carriers such as vegetable oils, dimethylacetamide, dimethylformamide, ethyl lactate, ethyl carbonate, isopropyl myristate, ethanol, polyols (glycerol, propylene glycol, liquid polyethylene glycol, and the like).

Aqueous injection suspensions can also contain substances that increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Non-lipid polycationic amino polymers can also be used for delivery. Optionally, the suspension can also contain suitable stabilizers or agents that increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Pharmaceutical compositions of the present invention can also be formulated to permit injectable, long-term, deposition. Injectable depot forms may be made by forming microencapsulated matrices of the compound in biodegradable polymers such as polylactide-polyglycolide. Depending upon the ratio of drug to polymer and the nature of the particular polymer employed, the rate of drug release can be controlled. Examples of other biodegradable polymers include poly(orthoesters) and poly(anhydrides). Depot injectable formulations are also prepared by entrapping the drug in microemulsions that are compatible with body tissues.

The pharmaceutical compositions of the present invention can be administered topically. For topical use the compounds of the present invention can also be prepared in suitable forms to be applied to the skin, or mucus membranes of the nose and throat, and can take the form of lotions, creams, ointments, liquid sprays or inhalants, drops, tinctures,

5 lozenges, or throat paints. Such topical formulations further can include chemical compounds such as dimethylsulfoxide (DMSO) to facilitate surface penetration of the active ingredient. In other transdermal formulations, typically in patch-delivered formulations, the pharmaceutically active compound is formulated with one or more skin penetrants, such as 2-N-methyl-pyrrolidone (NMP) or Azone. A topical semi-solid

10 ointment formulation typically contains a concentration of the active ingredient from about 1 to 20%, e.g., 5 to 10%, in a carrier such as a pharmaceutical cream base.

For application to the eyes or ears, the compounds of the present invention can be presented in liquid or semi-liquid form formulated in hydrophobic or hydrophilic bases as ointments, creams, lotions, paints or powders.

15 For rectal administration the compounds of the present invention can be administered in the form of suppositories admixed with conventional carriers such as cocoa butter, wax or other glyceride.

Inhalation formulations can also readily be formulated. For inhalation, various powder and liquid formulations can be prepared. For aerosol preparations, a sterile

20 formulation of the compound or salt form of the compound may be used in inhalers, such as metered dose inhalers, and nebulizers. Aerosolized forms may be especially useful for treating respiratory disorders.

Alternatively, the compounds of the present invention can be in powder form for reconstitution in the appropriate pharmaceutically acceptable carrier at the time of

25 delivery.

The pharmaceutically active compound in the pharmaceutical compositions of the present invention can be provided as the salt of a variety of acids, including but not limited to hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base

30 forms.

After pharmaceutical compositions have been prepared, they are packaged in an appropriate container and labeled for treatment of an indicated condition.

The active compound will be present in an amount effective to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

A "therapeutically effective dose" refers to that amount of active ingredient, for example OSP polypeptide, fusion protein, or fragments thereof, antibodies specific for OSP, agonists, antagonists or inhibitors of OSP, which ameliorates the signs or symptoms of the disease or prevent progression thereof; as would be understood in the medical arts, cure, although desired, is not required.

The therapeutically effective dose of the pharmaceutical agents of the present invention can be estimated initially by *in vitro* tests, such as cell culture assays, followed by assay in model animals, usually mice, rats, rabbits, dogs, or pigs. The animal model can also be used to determine an initial preferred concentration range and route of administration.

For example, the ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population) can be determined in one or more cell culture of animal model systems. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as LD50/ED50. Pharmaceutical compositions that exhibit large therapeutic indices are preferred.

The data obtained from cell culture assays and animal studies are used in formulating an initial dosage range for human use, and preferably provide a range of circulating concentrations that includes the ED50 with little or no toxicity. After administration, or between successive administrations, the circulating concentration of active agent varies within this range depending upon pharmacokinetic factors well known in the art, such as the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors specific to the subject requiring treatment. Factors that can be taken into account by the practitioner include the severity of the disease state, general health of the subject, age, weight, gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Where the therapeutic agent is a protein or antibody of the present invention, the therapeutic protein or antibody agent typically is administered at a daily dosage of 0.01 mg to 30 mg/kg of body weight of

5   the patient (*e.g.*, 1mg/kg to 5 mg/kg). The pharmaceutical formulation can be administered in multiple doses per day, if desired, to achieve the total desired daily dose.

Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors.

10   Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

Conventional methods, known to those of ordinary skill in the art of medicine, can be used to administer the pharmaceutical formulation(s) of the present invention to the patient. The pharmaceutical compositions of the present invention can be administered

15   alone, or in combination with other therapeutic agents or interventions.

Therapeutic Methods

The present invention further provides methods of treating subjects having defects in a gene of the invention, *e.g.*, in expression, activity, distribution, localization, and/or solubility, which can manifest as a disorder of ovarian function. As used herein,

20   "treating" includes all medically-acceptable types of therapeutic intervention, including palliation and prophylaxis (prevention) of disease. The term "treating" encompasses any improvement of a disease, including minor improvements. These methods are discussed below.

Gene Therapy and Vaccines

25   The isolated nucleic acids of the present invention can also be used to drive *in vivo* expression of the polypeptides of the present invention. *In vivo* expression can be driven from a vector, typically a viral vector, often a vector based upon a replication incompetent retrovirus, an adenovirus, or an adeno-associated virus (AAV), for the purpose of gene therapy. *In vivo* expression can also be driven from signals endogenous to the nucleic acid

30   or from a vector, often a plasmid vector, such as pVAX1 (Invitrogen, Carlsbad, CA, USA), for purpose of "naked" nucleic acid vaccination, as further described in U.S. Patent Nos. 5,589,466; 5,679,647; 5,804,566; 5,830,877; 5,843,913; 5,880,104; 5,958,891;

5,985,847; 6,017,897; 6,110,898; 6,204,250, the disclosures of which are incorporated herein by reference in their entireties. For cancer therapy, it is preferred that the vector also be tumor-selective. *See, e.g.*, Doronin *et al.*, *J. Virol.* 75: 3314-24 (2001).

    In another embodiment of the therapeutic methods of the present invention, a

5   therapeutically effective amount of a pharmaceutical composition comprising a nucleic acid molecule of the present invention is administered. The nucleic acid molecule can be delivered in a vector that drives expression of a OSP, fusion protein, or fragment thereof, or without such vector. Nucleic acid compositions that can drive expression of a OSP are administered, for example, to complement a deficiency in the native OSP, or as DNA

10   vaccines. Expression vectors derived from virus, replication deficient retroviruses, adenovirus, adeno-associated (AAV) virus, herpes virus, or vaccinia virus can be used as can plasmids. *See, e.g.*, Cid-Arregui, *supra*. In a preferred embodiment, the nucleic acid molecule encodes a OSP having the amino acid sequence of SEQ ID NO: 100-192, or a fragment, fusion protein, allelic variant or homolog thereof.

15   In still other therapeutic methods of the present invention, pharmaceutical compositions comprising host cells that express a OSP, fusions, or fragments thereof can be administered. In such cases, the cells are typically autologous, so as to circumvent xenogeneic or allotypic rejection, and are administered to complement defects in OSP production or activity. In a preferred embodiment, the nucleic acid molecules in the cells

20   encode a OSP having the amino acid sequence of SEQ ID NO: 100-192, or a fragment, fusion protein, allelic variant or homolog thereof.

Antisense Administration

    Antisense nucleic acid compositions, or vectors that drive expression of a OSG antisense nucleic acid, are administered to downregulate transcription and/or translation of

25   a OSG in circumstances in which excessive production, or production of aberrant protein, is the pathophysiologic basis of disease.

    Antisense compositions useful in therapy can have a sequence that is complementary to coding or to noncoding regions of a OSG. For example, oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10

30   and +10 from the start site, are preferred.

    Catalytic antisense compositions, such as ribozymes, that are capable of sequence-specific hybridization to OSG transcripts, are also useful in therapy. *See, e.g.*,

Phylactou, *Adv. Drug Deliv. Rev.* 44(2-3): 97-108 (2000); Phylactou *et al.*, *Hum. Mol. Genet.* 7(10): 1649-53 (1998); Rossi, *Ciba Found. Symp.* 209: 195-204 (1997); and Sigurdsson *et al.*, *Trends Biotechnol.* 13(8): 286-9 (1995).

5   Other nucleic acids useful in the therapeutic methods of the present invention are those that are capable of triplex helix formation in or near the OSG genomic locus. Such triplexing oligonucleotides are able to inhibit transcription. *See, e.g.,* Intody *et al.*, *Nucleic Acids Res.* 28(21): 4283-90 (2000); and McGuffie *et al.*, *Cancer Res.* 60(14): 3790-9 (2000). Pharmaceutical compositions comprising such triplex forming oligos (TFOs) are administered in circumstances in which excessive production, or production of aberrant

10   protein, is a pathophysiologic basis of disease.

In a preferred embodiment, the antisense molecule is derived from a nucleic acid molecule encoding a OSP, preferably a OSP comprising an amino acid sequence of SEQ ID NO: 100-192, or a fragment, allelic variant or homolog thereof. In a more preferred embodiment, the antisense molecule is derived from a nucleic acid molecule having a

15   nucleotide sequence of SEQ ID NO: 1-99, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

Polypeptide Administration

In one embodiment of the therapeutic methods of the present invention, a therapeutically effective amount of a pharmaceutical composition comprising a OSP, a

20   fusion protein, fragment, analog or derivative thereof is administered to a subject with a clinically-significant OSP defect.

Protein compositions are administered, for example, to complement a deficiency in native OSP. In other embodiments, protein compositions are administered as a vaccine to elicit a humoral and/or cellular immune response to OSP. The immune response can be

25   used to modulate activity of OSP or, depending on the immunogen, to immunize against aberrant or aberrantly expressed forms, such as mutant or inappropriately expressed isoforms. In yet other embodiments, protein fusions having a toxic moiety are administered to ablate cells that aberrantly accumulate OSP.

In a preferred embodiment, the polypeptide administered is a OSP comprising an

30   amino acid sequence of SEQ ID NO: 100-192, or a fusion protein, allelic variant, homolog, analog or derivative thereof. In a more preferred embodiment, the polypeptide

is encoded by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-99, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

Antibody, Agonist and Antagonist Administration

In another embodiment of the therapeutic methods of the present invention, a

5   therapeutically effective amount of a pharmaceutical composition comprising an antibody (including fragment or derivative thereof) of the present invention is administered. As is well known, antibody compositions are administered, for example, to antagonize activity of OSP, or to target therapeutic agents to sites of OSP presence and/or accumulation. In a preferred embodiment, the antibody specifically binds to a OSP comprising an amino acid

10 sequence of SEQ ID NO: 100-192, or a fusion protein, allelic variant, homolog, analog or derivative thereof. In a more preferred embodiment, the antibody specifically binds to a OSP encoded by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-99, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

The present invention also provides methods for identifying modulators which

15 bind to a OSP or have a modulatory effect on the expression or activity of a OSP. Modulators which decrease the expression or activity of OSP (antagonists) are believed to be useful in treating ovarian cancer. Such screening assays are known to those of skill in the art and include, without limitation, cell-based assays and cell-free assays. Small molecules predicted via computer imaging to specifically bind to regions of a OSP can

20 also be designed, synthesized and tested for use in the imaging and treatment of ovarian cancer. Further, libraries of molecules can be screened for potential anticancer agents by assessing the ability of the molecule to bind to the OSPs identified herein. Molecules identified in the library as being capable of binding to a OSP are key candidates for further evaluation for use in the treatment of ovarian cancer. In a preferred embodiment, these

25 molecules will downregulate expression and/or activity of a OSP in cells.

In another embodiment of the therapeutic methods of the present invention, a pharmaceutical composition comprising a non-antibody antagonist of OSP is administered. Antagonists of OSP can be produced using methods generally known in the art. In particular, purified OSP can be used to screen libraries of pharmaceutical agents,

30 often combinatorial libraries of small molecules, to identify those that specifically bind and antagonize at least one activity of a OSP.

In other embodiments a pharmaceutical composition comprising an agonist of a OSP is administered. Agonists can be identified using methods analogous to those used to identify antagonists.

In a preferred embodiment, the antagonist or agonist specifically binds to and antagonizes or agonizes, respectively, a OSP comprising an amino acid sequence of SEQ ID NO: 100-192, or a fusion protein, allelic variant, homolog, analog or derivative thereof. In a more preferred embodiment, the antagonist or agonist specifically binds to and antagonizes or agonizes, respectively, a OSP encoded by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-99, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

Targeting ovarian Tissue

The invention also provides a method in which a polypeptide of the invention, or an antibody thereto, is linked to a therapeutic agent such that it can be delivered to the ovarian or to specific cells in the ovarian. In a preferred embodiment, an anti-OSP antibody is linked to a therapeutic agent and is administered to a patient in need of such therapeutic agent. The therapeutic agent may be a toxin, if ovarian tissue needs to be selectively destroyed. This would be useful for targeting and killing ovarian cancer cells. In another embodiment, the therapeutic agent may be a growth or differentiation factor, which would be useful for promoting ovarian cell function.

In another embodiment, an anti-OSP antibody may be linked to an imaging agent that can be detected using, *e.g.*, magnetic resonance imaging, CT or PET. This would be useful for determining and monitoring ovarian function, identifying ovarian cancer tumors, and identifying noncancerous ovarian diseases.

EXAMPLES

Example 1: Gene Expression Analysis

Custom Microarray Experiment—Ovarian Cancer

Custom oligonucleotide microarrays were provided by Agilent Technologies, Inc. (Palo Alto, CA). The microarrays were fabricated by Agilent using their technology for the *in-situ* synthesis of 60mer oligonucleotides (Hughes, et al. 2001, Nature Biotechnology 19:342-347). The 60mer microarray probes were designed by Agilent, from gene

sequences provided by diaDexus, using Agilent proprietary algorithms. Whenever possible two different 60mers were designed for each gene of interest.

All microarray experiments were two-color experiments and were preformed using Agilent-recommended protocols and reagents. Briefly, each microarray was hybridized
5   with cRNAs synthesized from total RNA, isolated from cancer and normal tissues, labeled with fluorescent dyes Cyanine3 and Cyanine5 (NEN Life Science Products, Inc., Boston, MA) using a linear amplification method (Agilent). In each experiment the experimental sample was total RNA isolated from cancer tissue from a single individual and the reference sample was a pool of total RNA isolated from normal tissues of the same organ
10   as the cancerous tissue (*i.e.* normal ovarian tissue in experiments with ovarian cancer samples). Hybridizations were carried out at 60°C, overnight using Agilent *in-situ* hybridization buffer. Following washing, arrays were scanned with a GenePix 4000B Microarray Scanner (Axon Instruments, Inc., Union City, CA). The resulting images were analyzed with GenePix Pro 3.0 Microarray Acquisition and Analysis Software (Axon).
15   Two different chip designs were evaluated with overlapping sets of a total of 19 samples, comparing the expression patterns of ovarian cancer derived total RNA to total RNA isolated from a pool of 9 normal ovarian tissues. For Chip 2, all 19 samples (14 invasive carcinomas, 5 low malignant potential samples were analyzed and for Chip 1, a subset of 17 of these samples (13 invasive carcinomas, 4 low malignant potential samples) were
20   assessed.

Data normalization and expression profiling were done with Expressionist software from GeneData Inc. (Daly City, CA/Basel, Switzerland). Gene expression analysis was performed using only experiments that meet certain quality criteria. The quality criteria that experiments must meet are a combination of evaluations performed by
25   the Expressionist software and evaluations performed manually using raw and normalized data. To evaluate raw data quality, detection limits (the mean signal for a replicated negative control + 2 Standard Deviations (SD)) for each channel were calculated. The detection limit is a measure of non-specific hybridization. Acceptable detection limits were defined for each dye (80 for Cy5 and 150 for Cy3). Arrays with poor detection
30   limits in one or both channels were not analyzed and the experiments were repeated. To evaluate normalized data quality, positive control elements included in the array were utilized. These array features should have a mean ratio of 1 (no differential expression). If these features have a mean ratio of greater than 1.5-fold up or down, the experiments

were not analyzed further and were repeated. In addition to traditional scatter plots demonstrating the distribution of signal in each experiment, the Expressionist software also has minimum thresholding criteria that employ user defined parameters to identify quality data. These thresholds include two distinct quality measurements: 1) minimum

5 area percentage, which is a measure of the integrity of each spot and 2) signal to noise ratio, which ensures that the signal being measured is significantly above any background (nonspecific) signal present. Only those features that meet the threshhold criteria were included in the filtering and analyses carried out by Expressionist. The thresholding settings employed require a minimum area percentage of 60% [(% pixels > background +

10 2SD)-(% pixels saturated)], and a minimum signal to noise ratio of 2.0 in both channels. By these criteria, very low expressors, saturated features and spots with abnormally high local background were not included in analysis.

Relative expression data was collected from Expressionist based on filtering and clustering analyses. Up-regulated genes were identified using criteria for the percentage

15 of experiments in which the gene is up-regulated. In general, up-regulation in ~30% of samples tested was used as a cutoff for filtering. The results for the statistically significant up-regulated genes on Chip 1 are shown in Table 1. The results for Chip 2 are shown in Table 2. The first two columns of each table contain information about the sequence itself (DEX ID, OligoID), the next columns show the results obtained for all ("ALL") ovarian

20 cancer samples, invasive carcinomas ("INV") and low malignant potential ("LMP") samples. '%up' indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed (n=17 for Chip 1, n=19 for Chip 2), '%valid up' indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

25 Table 1. Expression data from Chip 1 for up-regulated genes in ovarian cancer samples.

DEX ID	OligoID	%up ALL n=17	%valid up ALL n=17	%up INV n=13	%valid up INV n=13	%up LMP n=4	%valid up LMP n=4
DEX0367_001.nt.1	23690.01	23.5	23.5	23.1	23.1	25	25
DEX0367_001.nt.1	23690.02	29.4	29.4	30.8	30.8	25	25
DEX0367_001.nt.2	23690.01	23.5	23.5	23.1	23.1	25	25
DEX0367_001.nt.2	23690.02	29.4	29.4	30.8	30.8	25	25
DEX0367_002.nt.1	34930.01	23.5	23.5	30.8	30.8	0	0
DEX0367_002.nt.1	34930.02	23.5	23.5	30.8	30.8	0	0
DEX0367_003.nt.1	96483.01	64.7	68.8	76.9	83.3	25	25
DEX0367_003.nt.1	96483.02	64.7	68.8	76.9	83.3	25	25
DEX0367_003.nt.1	103465.01	64.7	68.8	76.9	83.3	25	25
DEX0367_003.nt.1	103465.02	58.8	76.9	69.2	90	25	33.3
DEX0367_003.nt.1	103469.01	70.6	80	84.6	91.7	25	33.3

DEX0367_003.nt.1	103469.02	58.8	90.9	69.2	100	25	50
DEX0367_004.nt.1	11373.01	41.2	41.2	38.5	38.5	50	50
DEX0367_004.nt.1	11373.02	41.2	41.2	38.5	38.5	50	50
DEX0367_004.nt.1	21559.01	47.1	47.1	46.2	46.2	50	50
DEX0367_004.nt.1	21559.02	47.1	47.1	46.2	46.2	50	50
DEX0367_007.nt.1	24874.02	29.4	31.2	23.1	25	50	50
DEX0367_007.nt.2	24874.02	29.4	31.2	23.1	25	50	50
DEX0367_008.nt.1	23378.01	11.8	18.2	15.4	20	0	0
DEX0367_008.nt.1	23378.02	17.6	23.1	7.7	9.1	50	100
DEX0367_009.nt.1	30109.01	41.2	46.7	30.8	33.3	75	100
DEX0367_009.nt.1	30109.02	35.3	40	23.1	27.3	75	75
DEX0367_010.nt.1	18508.01	23.5	44.4	30.8	44.4	0	0
DEX0367_010.nt.1	18508.02	17.6	23.1	23.1	30	0	0
DEX0367_010.nt.1	22387.01	35.3	54.5	46.2	66.7	0	0
DEX0367_010.nt.1	22387.02	41.2	43.8	53.8	58.3	0	0
DEX0367_011.nt.1	9720.01	47.1	47.1	38.5	38.5	75	75
DEX0367_011.nt.1	9720.02	52.9	52.9	46.2	46.2	75	75
DEX0367_012.nt.1	20681.01	35.3	35.3	38.5	38.5	25	25
DEX0367_012.nt.1	20681.02	29.4	31.2	38.5	38.5	0	0
DEX0367_012.nt.2	20681.01	35.3	35.3	38.5	38.5	25	25
DEX0367_012.nt.2	20681.02	29.4	31.2	38.5	38.5	0	0
DEX0367_015.nt.1	9838.01	35.3	42.9	38.5	45.5	25	33.3
DEX0367_015.nt.1	9838.02	35.3	37.5	38.5	38.5	25	33.3
DEX0367_015.nt.2	9838.01	35.3	42.9	38.5	45.5	25	33.3
DEX0367_015.nt.2	9838.02	35.3	37.5	38.5	38.5	25	33.3
DEX0367_016.nt.1	10624.01	17.6	50	15.4	50	25	50
DEX0367_016.nt.1	10624.02	17.6	50	7.7	33.3	50	66.7
DEX0367_016.nt.2	10624.01	17.6	50	15.4	50	25	50
DEX0367_016.nt.2	10624.02	17.6	50	7.7	33.3	50	66.7
DEX0367_017.nt.1	19518.01	29.4	29.4	38.5	38.5	0	0
DEX0367_017.nt.1	19518.02	29.4	29.4	38.5	38.5	0	0
DEX0367_019.nt.1	16063.01	23.5	23.5	7.7	7.7	75	75
DEX0367_019.nt.1	16063.02	23.5	23.5	7.7	7.7	75	75
DEX0367_020.nt.1	22573.01	52.9	52.9	69.2	69.2	0	0
DEX0367_020.nt.1	22573.02	52.9	52.9	69.2	69.2	0	0
DEX0367_021.nt.1	21575.01	41.2	41.2	46.2	46.2	25	25
DEX0367_021.nt.1	21575.02	41.2	41.2	46.2	46.2	25	25
DEX0367_021.nt.2	21575.01	41.2	41.2	46.2	46.2	25	25
DEX0367_021.nt.2	21575.02	41.2	41.2	46.2	46.2	25	25
DEX0367_022.nt.1	20669.01	35.3	42.9	46.2	50	0	0
DEX0367_022.nt.1	20669.02	35.3	46.2	46.2	50	0	0
DEX0367_023.nt.1	11469.01	47.1	47.1	61.5	61.5	0	0
DEX0367_023.nt.1	11469.02	52.9	52.9	61.5	61.5	25	25
DEX0367_023.nt.2	11469.01	47.1	47.1	61.5	61.5	0	0
DEX0367_023.nt.2	11469.02	52.9	52.9	61.5	61.5	25	25
DEX0367_024.nt.1	21433.01	64.7	64.7	61.5	61.5	75	75
DEX0367_024.nt.1	21433.02	64.7	64.7	61.5	61.5	75	75
DEX0367_024.nt.1	21469.01	70.6	70.6	61.5	61.5	100	100
DEX0367_024.nt.1	21469.02	82.4	82.4	76.9	76.9	100	100
DEX0367_024.nt.1	21475.01	58.8	58.8	53.8	53.8	75	75
DEX0367_024.nt.1	21475.02	52.9	52.9	53.8	53.8	50	50
DEX0367_024.nt.1	23780.01	47.1	47.1	46.2	46.2	50	50
DEX0367_024.nt.1	23780.02	41.2	50	46.2	54.5	25	33.3
DEX0367_024.nt.2	21433.01	64.7	64.7	61.5	61.5	75	75
DEX0367_024.nt.2	21433.02	64.7	64.7	61.5	61.5	75	75
DEX0367_024.nt.2	21469.01	70.6	70.6	61.5	61.5	100	100
DEX0367_024.nt.2	21469.02	82.4	82.4	76.9	76.9	100	100

DEX0367_024.nt.2	21475.01	58.8	58.8	53.8	53.8	75	75
DEX0367_024.nt.2	21475.02	52.9	52.9	53.8	53.8	50	50
DEX0367_024.nt.2	23780.01	47.1	47.1	46.2	46.2	50	50
DEX0367_024.nt.2	23780.02	41.2	50	46.2	54.5	25	33.3
DEX0367_024.nt.3	21433.01	64.7	64.7	61.5	61.5	75	75
DEX0367_024.nt.3	21433.02	64.7	64.7	61.5	61.5	75	75
DEX0367_024.nt.3	21469.01	70.6	70.6	61.5	61.5	100	100
DEX0367_024.nt.3	21469.02	82.4	82.4	76.9	76.9	100	100
DEX0367_024.nt.3	21475.01	58.8	58.8	53.8	53.8	75	75
DEX0367_024.nt.3	21475.02	52.9	52.9	53.8	53.8	50	50
DEX0367_024.nt.3	23780.01	47.1	47.1	46.2	46.2	50	50
DEX0367_024.nt.3	23780.02	41.2	50	46.2	54.5	25	33.3
DEX0367_024.nt.4	21433.01	64.7	64.7	61.5	61.5	75	75
DEX0367_024.nt.4	21433.02	64.7	64.7	61.5	61.5	75	75
DEX0367_024.nt.4	21469.01	70.6	70.6	61.5	61.5	100	100
DEX0367_024.nt.4	21469.02	82.4	82.4	76.9	76.9	100	100
DEX0367_024.nt.4	21475.01	58.8	58.8	53.8	53.8	75	75
DEX0367_024.nt.4	21475.02	52.9	52.9	53.8	53.8	50	50
DEX0367_024.nt.4	23780.01	47.1	47.1	46.2	46.2	50	50
DEX0367_024.nt.4	23780.02	41.2	50	46.2	54.5	25	33.3
DEX0367_025.nt.1	9920.01	23.5	23.5	30.8	30.8	0	0
DEX0367_025.nt.1	9920.02	23.5	23.5	30.8	30.8	0	0
DEX0367_025.nt.2	9920.01	23.5	23.5	30.8	30.8	0	0
DEX0367_025.nt.2	9920.02	23.5	23.5	30.8	30.8	0	0
DEX0367_025.nt.3	9920.01	23.5	23.5	30.8	30.8	0	0
DEX0367_025.nt.3	9920.02	23.5	23.5	30.8	30.8	0	0
DEX0367_026.nt.1	91971.01	94.1	94.1	92.3	92.3	100	100
DEX0367_026.nt.1	91971.02	94.1	94.1	92.3	92.3	100	100
DEX0367_027.nt.1	16374.01	23.5	23.5	30.8	30.8	0	0
DEX0367_027.nt.1	16374.02	17.6	17.6	23.1	23.1	0	0
DEX0367_028.nt.1	12149.01	52.9	52.9	46.2	46.2	75	75
DEX0367_028.nt.1	12149.02	47.1	47.1	38.5	38.5	75	75
DEX0367_028.nt.1	21507.01	29.4	29.4	23.1	23.1	50	50
DEX0367_028.nt.1	21507.02	29.4	31.2	23.1	25	50	50
DEX0367_028.nt.1	21547.01	47.1	47.1	46.2	46.2	50	50
DEX0367_028.nt.1	21547.02	41.2	41.2	38.5	38.5	50	50
DEX0367_028.nt.2	12149.01	52.9	52.9	46.2	46.2	75	75
DEX0367_028.nt.2	12149.02	47.1	47.1	38.5	38.5	75	75
DEX0367_028.nt.2	21507.01	29.4	29.4	23.1	23.1	50	50
DEX0367_028.nt.2	21507.02	29.4	31.2	23.1	25	50	50
DEX0367_028.nt.2	21547.01	47.1	47.1	46.2	46.2	50	50
DEX0367_028.nt.2	21547.02	41.2	41.2	38.5	38.5	50	50
DEX0367_030.nt.1	16063.01	23.5	23.5	7.7	7.7	75	75
DEX0367_030.nt.1	16063.02	23.5	23.5	7.7	7.7	75	75
DEX0367_031.nt.1	21549.01	29.4	31.2	23.1	25	50	50
DEX0367_031.nt.1	21549.02	29.4	31.2	23.1	25	50	50
DEX0367_032.nt.1	23466.01	17.6	20	23.1	27.3	0	0
DEX0367_032.nt.1	23466.02	23.5	23.5	30.8	30.8	0	0
DEX0367_033.nt.1	17430.01	23.5	23.5	30.8	30.8	0	0
DEX0367_033.nt.1	17430.02	17.6	17.6	23.1	23.1	0	0
DEX0367_033.nt.1	23466.01	17.6	20	23.1	27.3	0	0
DEX0367_033.nt.1	23466.02	23.5	23.5	30.8	30.8	0	0
DEX0367_033.nt.2	17430.01	23.5	23.5	30.8	30.8	0	0
DEX0367_033.nt.2	17430.02	17.6	17.6	23.1	23.1	0	0
DEX0367_033.nt.2	23466.01	17.6	20	23.1	27.3	0	0
DEX0367_033.nt.2	23466.02	23.5	23.5	30.8	30.8	0	0
DEX0367_034.nt.1	17278.01	23.5	23.5	30.8	30.8	0	0

DEX0367_034.nt.1	17278.02	17.6	17.6	23.1	23.1	0	0
DEX0367_034.nt.2	17278.01	23.5	23.5	30.8	30.8	0	0
DEX0367_034.nt.2	17278.02	17.6	17.6	23.1	23.1	0	0
DEX0367_035.nt.1	20773.01	17.6	17.6	23.1	23.1	0	0
DEX0367_035.nt.1	20773.02	23.5	25	30.8	33.3	0	0
DEX0367_035.nt.2	20773.01	17.6	17.6	23.1	23.1	0	0
DEX0367_035.nt.2	20773.02	23.5	25	30.8	33.3	0	0
DEX0367_035.nt.3	20773.01	17.6	17.6	23.1	23.1	0	0
DEX0367_035.nt.3	20773.02	23.5	25	30.8	33.3	0	0
DEX0367_036.nt.1	18520.01	23.5	23.5	30.8	30.8	0	0
DEX0367_036.nt.1	18520.02	17.6	17.6	23.1	23.1	0	0
DEX0367_036.nt.1	22734.01	5.9	5.9	7.7	7.7	0	0
DEX0367_036.nt.1	22734.02	23.5	23.5	30.8	30.8	0	0
DEX0367_038.nt.1	10722.01	82.4	82.4	84.6	84.6	75	75
DEX0367_038.nt.1	10722.02	76.5	81.2	84.6	84.6	50	66.7
DEX0367_039.nt.1	103385.01	58.8	58.8	76.9	76.9	0	0
DEX0367_039.nt.1	103385.02	58.8	58.8	76.9	76.9	0	0
DEX0367_039.nt.2	103385.01	58.8	58.8	76.9	76.9	0	0
DEX0367_039.nt.2	103385.02	58.8	58.8	76.9	76.9	0	0
DEX0367_039.nt.3	103385.01	58.8	58.8	76.9	76.9	0	0
DEX0367_039.nt.3	103385.02	58.8	58.8	76.9	76.9	0	0
DEX0367_040.nt.1	92327.01	52.9	56.2	61.5	66.7	25	25
DEX0367_040.nt.1	92327.02	52.9	52.9	61.5	61.5	25	25
DEX0367_040.nt.2	92327.01	52.9	56.2	61.5	66.7	25	25
DEX0367_040.nt.2	92327.02	52.9	52.9	61.5	61.5	25	25
DEX0367_040.nt.3	92327.01	52.9	56.2	61.5	66.7	25	25
DEX0367_040.nt.3	92327.02	52.9	52.9	61.5	61.5	25	25
DEX0367_040.nt.4	92327.01	52.9	56.2	61.5	66.7	25	25
DEX0367_040.nt.4	92327.02	52.9	52.9	61.5	61.5	25	25
DEX0367_041.nt.1	11575.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.1	11575.02	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.1	17486.01	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.1	17486.02	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.1	17490.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.1	17490.02	58.8	58.8	53.8	53.8	75	75
DEX0367_041.nt.2	11575.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.2	11575.02	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.2	17486.01	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.2	17486.02	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.2	17490.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.2	17490.02	58.8	58.8	53.8	53.8	75	75
DEX0367_041.nt.3	11575.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.3	11575.02	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.3	17486.01	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.3	17486.02	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.3	17490.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.3	17490.02	58.8	58.8	53.8	53.8	75	75
DEX0367_041.nt.4	11575.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.4	11575.02	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.4	17486.01	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.4	17486.02	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.4	17490.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.4	17490.02	58.8	58.8	53.8	53.8	75	75
DEX0367_041.nt.5	11575.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.5	11575.02	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.5	17486.01	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.5	17486.02	47.1	47.1	46.2	46.2	50	50

DEX0367_041.nt.5	17490.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.5	17490.02	58.8	58.8	53.8	53.8	75	75
DEX0367_041.nt.6	11575.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.6	11575.02	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.6	17486.01	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.6	17486.02	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.6	17490.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.6	17490.02	58.8	58.8	53.8	53.8	75	75
DEX0367_041.nt.7	11575.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.ht.7	11575.02	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.7	17486.01	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.7	17486.02	47.1	47.1	46.2	46.2	50	50
DEX0367_041.nt.7	17490.01	52.9	52.9	53.8	53.8	50	50
DEX0367_041.nt.7	17490.02	58.8	58.8	53.8	53.8	75	75
DEX0367_043.nt.1	21505.01	94.1	94.1	92.3	92.3	100	100
DEX0367_043.nt.1	21505.02	94.1	94.1	92.3	92.3	100	100
DEX0367_043.nt.2	11527.01	88.2	88.2	84.6	84.6	100	100
DEX0367_043.nt.2	11527.02	88.2	88.2	84.6	84.6	100	100
DEX0367_044.nt.1	21501.01	47.1	50	61.5	61.5	0	0
DEX0367_044.nt.1	21501.02	41.2	41.2	53.8	53.8	0	0
DEX0367_044.nt.1	21511.01	47.1	47.1	61.5	61.5	0	0
DEX0367_044.nt.1	21511.02	47.1	47.1	53.8	53.8	25	25
DEX0367_044.nt.2	21501.01	47.1	50	61.5	61.5	0	0
DEX0367_044.nt.2	21501.02	41.2	41.2	53.8	53.8	0	0
DEX0367_044.nt.2	21511.01	47.1	47.1	61.5	61.5	0	0
DEX0367_044.nt.2	21511.02	47.1	47.1	53.8	53.8	25	25
DEX0367_045.nt.1	12155.01	23.5	23.5	30.8	30.8	0	0
DEX0367_045.nt.1	12155.02	23.5	23.5	30.8	30.8	0	0
DEX0367_045.nt.2	12155.01	23.5	23.5	30.8	30.8	0	0
DEX0367_045.nt.2	12155.02	23.5	23.5	30.8	30.8	0	0
DEX0367_046.nt.1	10702.01	94.1	94.1	92.3	92.3	100	100
DEX0367_046.nt.1	10702.02	94.1	94.1	92.3	92.3	100	100
DEX0367_048.nt.1	9958.01	47.1	47.1	46.2	46.2	50	50
DEX0367_048.nt.1	9958.02	47.1	47.1	46.2	46.2	50	50
DEX0367_049.nt.1	36013.01	23.5	23.5	7.7	7.7	75	75
DEX0367_049.nt.1	36013.02	11.8	11.8	0	0	50	50
DEX0367_050.nt.1	17314.01	23.5	26.7	15.4	16.7	50	66.7
DEX0367_050.nt.1	17314.02	23.5	26.7	15.4	16.7	50	66.7
DEX0367_051.nt.1	22573.01	52.9	52.9	69.2	69.2	0	0
DEX0367_051.nt.1	22573.02	52.9	52.9	69.2	69.2	0	0
DEX0367_054.nt.1	11511.01	94.1	100	92.3	100	100	100
DEX0367_054.nt.1	11511.02	88.2	100	84.6	100	100	100
DEX0367_055.nt.1	23378.01	11.8	18.2	15.4	20	0	0
DEX0367_055.nt.1	23378.02	17.6	23.1	7.7	9.1	50	100

Table 2. Expression data from Chip 2 for up-regulated genes in ovarian cancer samples.

DEX ID	OligoID	%up ALL n=19	%valid up ALL n=19	%up INV n=14	%valid up INV n=14	%up LMP n=5	%valid up LMP n=5
DEX0367_013.nt.1	35316	26.3	33.3	7.1	9.1	80	100
DEX0367_013.nt.1	35316.1	31.6	42.9	14.3	20	80	100
DEX0367_013.nt.1	35317	31.6	42.9	14.3	20	80	100
DEX0367_013.nt.1	35317.1	31.6	35.3	14.3	15.4	80	100
DEX0367_013.nt.2	35316	26.3	33.3	7.1	9.1	80	100
DEX0367_013.nt.2	35316.1	31.6	42.9	14.3	20	80	100
DEX0367_013.nt.2	35317	31.6	42.9	14.3	20	80	100
DEX0367_013.nt.2	35317.1	31.6	35.3	14.3	15.4	80	100

DEX0367_014.nt.1	34334	89.5	89.5	85.7	85.7	100	100
DEX0367_014.nt.1	34334.1	84.2	88.9	85.7	85.7	80	.100
DEX0367_014.nt.1	34335	94.7	100	92.9	100	100	100
DEX0367_014.nt.1	34335.1	89.5	100	92.9	100	80	100
DEX0367_014.nt.2	34334	89.5	89.5	85.7	85.7	100	100
DEX0367_014.nt.2	34334.1	84.2	88.9	85.7	85.7	80	100
DEX0367_014.nt.2	34335	94.7	100	92.9	100	100	.100
DEX0367_014.nt.2	34335.1	89.5	100	92.9	100	80	100
DEX0367_019.nt.1	36482	31.6	42.9	28.6	40	40	50
DEX0367_019.nt.1	36482.1	31.6	50	28.6	44.4	40	66.7
DEX0367_019.nt.2	36482	31.6	42.9	28.6	40	40	50
DEX0367_019.nt.2	36482.1	31.6	50	28.6	44.4	40	66.7
DEX0367_037.nt.1	2023	21.1	23.5	28.6	30.8	0	0
DEX0367_037.nt.1	5327*	21.1	26.7	28.6	30.8	0	0
DEX0367_042.nt.1	23543*	31.6	35.3	35.7	38.5	20	25
DEX0367_042.nt.1	23543.1*	26.3	26.3	28.6	28.6	20	20
DEX0367_042.nt.2	23542*	15.8	16.7	21.4	23.1	0	0
DEX0367_042.nt.2	23542.1*	26.3	26.3	35.7	35.7	0	0
DEX0367_042.nt.2	23543*	31.6	35.3	35.7	38.5	20	.25
DEX0367_042.nt.2	23543.1*	26.3	26.3	28.6	28.6	20	.20
DEX0367_042.nt.3	23542*	15.8	16.7	21.4	23.1	0	0
DEX0367_042.nt.3	23542.1*	26.3	26.3	35.7	35.7	0	0
DEX0367_042.nt.3	23543*	31.6	35.3	35.7	38.5	20	.25
DEX0367_042.nt.3	23543.1*	26.3	26.3	28.6	28.6	20	20

*data from lower PMT voltage scan

The location/mapping of the oligos on the ovarian Cancer genes were as follows:

SeqID	Master OligoID	oligoID	ChipName	Location
DEX0367_001.nt.1	23690	23690.01	Chip 1	575-631
DEX0367_001.nt.1	23690	23690.02	Chip 1	575-631
DEX0367_001.nt.2	23690	23690.01	Chip 1	905-964
DEX0367_001.nt.2	23690	23690.02	Chip 1	905-964
DEX0367_002.nt.1	34930	34930.01	Chip 1	4736-4795
DEX0367_002.nt.1	34930	34930.02	Chip 1	4736-4795
DEX0367_003.nt.1	103465	103465.01	Chip 1	865-924
DEX0367_003.nt.1	103465	103465.02	Chip 1	865-924
DEX0367_003.nt.1	103469	103469.01	Chip 1	857-916
DEX0367_003.nt.1	103469	103469.02	Chip 1	857-916
DEX0367_003.nt.1	96483	96483.01	Chip 1	865-924
DEX0367_003.nt.1	96483	96483.02	Chip 1	865-924
DEX0367_004.nt.1	11373	11373.01	Chip 1	1442-1494
DEX0367_004.nt.1	11373	11373.02	Chip 1	1442-1494
DEX0367_004.nt.1	21559	21559.01	Chip 1	1437-1492
DEX0367_004.nt.1	21559	21559.02	Chip 1	1437-1492
DEX0367_005.nt.1	17466	17466.01	Chip 1	1075-1134
DEX0367_005.nt.1	17466	17466.02	Chip 1	1075-1134
DEX0367_006.nt.1	105996	105996	Chip 2	8462-8521
DEX0367_006.nt.1	105996	105996.1	Chip 2	8462-8521
DEX0367_006.nt.2	105996	105996	Chip 2	7306-7365
DEX0367_006.nt.2	105996	105996.1	Chip 2	7306-7365
DEX0367_007.nt.1	24874	24874.01	Chip 1	475-534
DEX0367_007.nt.1	24874	24874.02	Chip 1	475-534
DEX0367_007.nt.2	24874	24874.01	Chip 1	475-534
DEX0367_007.nt.2	24874	24874.02	Chip 1	475-534
DEX0367_008.nt.1	23378	23378	Chip 2	69-15
DEX0367_008.nt.1	23378	23378.01	Chip 1	69-15

DEX0367_008.nt.1	23378	23378.02	Chip 1	69-15
DEX0367_008.nt.1	23378	23378.1	Chip 2	69-15
DEX0367_009.nt.1	30109	30109.01	Chip 1	982-1041
DEX0367_009.nt.1	30109	30109.02	Chip 1	982-1041
DEX0367_010.nt.1	18508	18508.01	Chip 1	1193-1252
DEX0367_010.nt.1	18508	18508.02	Chip 1	1193-1252
DEX0367_010.nt.1	22387	22387.01	Chip 1	1666-1725
DEX0367_010.nt.1	22387	22387.02	Chip 1	1666-1725
DEX0367_011.nt.1	9720	9720.01	Chip 1	1745-1804
DEX0367_011.nt.1	9720	9720.02	Chip 1	1745-1804
DEX0367_012.nt.1	20681	20681.01	Chip 1	826-881
DEX0367_012.nt.1	20681	20681.02	Chip 1	826-881
DEX0367_012.nt.2	20681	20681.01	Chip 1	663-718
DEX0367_012.nt.2	20681	20681.02	Chip 1	663-718
DEX0367_013.nt.1	35316	35316	Chip 2	688-742
DEX0367_013.nt.1	35316	35316.1	Chip 2	688-742
DEX0367_013.nt.1	35317	35317	Chip 2	668-727
DEX0367_013.nt.1	35317	35317.1	Chip 2	668-727
DEX0367_013.nt.2	35316	35316	Chip 2	418-472
DEX0367_013.nt.2	35316	35316.1	Chip 2	418-472
DEX0367_013.nt.2	35317	35317	Chip 2	398-457
DEX0367_013.nt.2	35317	35317.1	Chip 2	398-457
DEX0367_014.nt.1	34334	34334	Chip 2	3108-3167
DEX0367_014.nt.1	34334	34334.1	Chip 2	3108-3167
DEX0367_014.nt.1	34335	34335	Chip 2	3022-3081
DEX0367_014.nt.1	34335	34335.1	Chip 2	3022-3081
DEX0367_014.nt.2	34334	34334	Chip 2	2527-2586
DEX0367_014.nt.2	34334	34334.1	Chip 2	2527-2586
DEX0367_014.nt.2	34335	34335	Chip 2	2441-2500
DEX0367_014.nt.2	34335	34335.1	Chip 2	2441-2500
DEX0367_015.nt.1	9838	9838.01	Chip 1	1304-1363
DEX0367_015.nt.1	9838	9838.02	Chip 1	1304-1363
DEX0367_015.nt.2	9838	9838.01	Chip 1	1277-1336
DEX0367_015.nt.2	9838	9838.02	Chip 1	1277-1336
DEX0367_016.nt.1	10624	10624.01	Chip 1	1832-1891
DEX0367_016.nt.1	10624	10624.02	Chip 1	1832-1891
DEX0367_016.nt.2	10624	10624.01	Chip 1	1155-1214
DEX0367_016.nt.2	10624	10624.02	Chip 1	1155-1214
DEX0367_017.nt.1	19518	19518.01	Chip 1	277-336
DEX0367_017.nt.1	19518	19518.02	Chip 1	277-336
DEX0367_018.nt.1	23734	23734.01	Chip 1	531-590
DEX0367_018.nt.1	23734	23734.02	Chip 1	531-590
DEX0367_019.nt.1	16063	16063.01	Chip 1	70-11
DEX0367_019.nt.1	16063	16063.02	Chip 1	70-11
DEX0367_019.nt.1	36482	36482	Chip 2	314-373
DEX0367_019.nt.1	36482	36482.1	Chip 2	314-373
DEX0367_019.nt.2	36482	36482	Chip 2	930-989
DEX0367_019.nt.2	36482	36482.1	Chip 2	930-989
DEX0367_020.nt.1	22573	22573.01	Chip 1	652-711
DEX0367_020.nt.1	22573	22573.02	Chip 1	652-711
DEX0367_021.nt.1	21575	21575.01	Chip 1	1516-1575
DEX0367_021.nt.1	21575	21575.02	Chip 1	1516-1575
DEX0367_021.nt.2	21575	21575.01	Chip 1	2287-2346
DEX0367_021.nt.2	21575	21575.02	Chip 1	2287-2346
DEX0367_022.nt.1	20669	20669.01	Chip 1	615-674
DEX0367_022.nt.1	20669	20669.02	Chip 1	615-674
DEX0367_023.nt.1	11469	11469.01	Chip 1	1210-1151

DEX0367_023.nt.1	11469	11469.02	Chip 1	1210-1151
DEX0367_023.nt.2	11469	11469.01	Chip 1	812-753
DEX0367_023.nt.2	11469	11469.02	Chip 1	812-753
DEX0367_024.nt.1	21433	21433.01	Chip 1	518-575
DEX0367_024.nt.1	21433	21433.02	Chip 1	518-575
DEX0367_024.nt.1	21469	21469.01	Chip 1	430-489
DEX0367_024.nt.1	21469	21469.02	Chip 1	430-489
DEX0367_024.nt.1	21475	21475.01	Chip 1	517-576
DEX0367_024.nt.1	21475	21475.02	Chip 1	517-576
DEX0367_024.nt.1	23780	23780.01	Chip 1	517-576
DEX0367_024.nt.1	23780	23780.02	Chip 1	517-576
DEX0367_024.nt.2	21433	21433.01	Chip 1	1616-1673
DEX0367_024.nt.2	21433	21433.02	Chip 1	1616-1673
DEX0367_024.nt.2	21469	21469.01	Chip 1	1528-1587
DEX0367_024.nt.2	21469	21469.02	Chip 1	1528-1587
DEX0367_024.nt.2	21475	21475.01	Chip 1	1615-1674
DEX0367_024.nt.2	21475	21475.02	Chip 1	1615-1674
DEX0367_024.nt.2	23780	23780.01	Chip 1	1615-1674
DEX0367_024.nt.2	23780	23780.02	Chip 1	1615-1674
DEX0367_024.nt.3	21433	21433.01	Chip 1	1859-1916
DEX0367_024.nt.3	21433	21433.02	Chip 1	1859-1916
DEX0367_024.nt.3	21469	21469.01	Chip 1	1771-1830
DEX0367_024.nt.3	21469	21469.02	Chip 1	1771-1830
DEX0367_024.nt.3	21475	21475.01	Chip 1	1858-1917
DEX0367_024.nt.3	21475	21475.02	Chip 1	1858-1917
DEX0367_024.nt.3	23780	23780.01	Chip 1	1858-1917
DEX0367_024.nt.3	23780	23780.02	Chip 1	1858-1917
DEX0367_024.nt.4	21433	21433.01	Chip 1	2002-2059
DEX0367_024.nt.4	21433	21433.02	Chip 1	2002-2059
DEX0367_024.nt.4	21469	21469.01	Chip 1	1914-1973
DEX0367_024.nt.4	21469	21469.02	Chip 1	1914-1973
DEX0367_024.nt.4	21475	21475.01	Chip 1	2001-2060
DEX0367_024.nt.4	21475	21475.02	Chip 1	2001-2060
DEX0367_024.nt.4	23780	23780.01	Chip 1	2001-2060
DEX0367_024.nt.4	23780	23780.02	Chip 1	2001-2060
DEX0367_025.nt.1	9920	9920.01	Chip 1	1022-1081
DEX0367_025.nt.1	9920	9920.02	Chip 1	1022-1081
DEX0367_025.nt.2	9920	9920.01	Chip 1	1016-1075
DEX0367_025.nt.2	9920	9920.02	Chip 1	1016-1075
DEX0367_025.nt.3	9920	9920.01	Chip 1	613-672
DEX0367_025.nt.3	9920	9920.02	Chip 1	613-672
DEX0367_026.nt.1	91971	91971.01	Chip 1	1686-1745
DEX0367_026.nt.1	91971	91971.02	Chip 1	1686-1745
DEX0367_027.nt.1	16374	16374.01	Chip 1	2119-2177
DEX0367_027.nt.1	16374	16374.02	Chip 1	2119-2177
DEX0367_028.nt.1	12149	12149.01	Chip 1	2439-2498
DEX0367_028.nt.1	12149	12149.02	Chip 1	2439-2498
DEX0367_028.nt.1	21507	21507.01	Chip 1	2357-2416
DEX0367_028.nt.1	21507	21507.02	Chip 1	2357-2416
DEX0367_028.nt.1	21547	21547.01	Chip 1	1555-1614
DEX0367_028.nt.1	21547	21547.02	Chip 1	1555-1614
DEX0367_028.nt.2	12149	12149.01	Chip 1	1872-1931
DEX0367_028.nt.2	12149	12149.02	Chip 1	1872-1931
DEX0367_028.nt.2	21507	21507.01	Chip 1	1790-1849
DEX0367_028.nt.2	21507	21507.02	Chip 1	1790-1849
DEX0367_028.nt.2	21547	21547.01	Chip 1	988-1047
DEX0367_028.nt.2	21547	21547.02	Chip 1	988-1047

DEX0367_029.nt.1	12167	12167.01	Chip 1	475-534
DEX0367_029.nt.1	12167	12167.02	Chip 1	475-534
DEX0367_029.nt.2	12167	12167.01	Chip 1	475-534
DEX0367_029.nt.2	12167	12167.02	Chip 1	475-534
DEX0367_029.nt.3	12167	12167.01	Chip 1	475-534
DEX0367_029.nt.3	12167	12167.02	Chip 1	475-534
DEX0367_029.nt.4	12167	12167.01	Chip 1	475-534
DEX0367_029.nt.4	12167	12167.02	Chip 1	475-534
DEX0367_030.nt.1	16063	16063.01	Chip 1	93-35
DEX0367_030.nt.1	16063	16063.02	Chip 1	93-35
DEX0367_031.nt.1	21549	21549.01	Chip 1	1483-1542
DEX0367_031.nt.1	21549	21549.02	Chip 1	1483-1542
DEX0367_032.nt.1	23466	23466.01	Chip 1	517-572
DEX0367_032.nt.1	23466	23466.02	Chip 1	517-572
DEX0367_033.nt.1	17430	17430.01	Chip 1	2388-2447
DEX0367_033.nt.1	17430	17430.02	Chip 1	2388-2447
DEX0367_033.nt.1	23466	23466.01	Chip 1	2024-2080
DEX0367_033.nt.1	23466	23466.02	Chip 1	2024-2080
DEX0367_033.nt.2	17430	17430.01	Chip 1	2816-2875
DEX0367_033.nt.2	17430	17430.02	Chip 1	2816-2875
DEX0367_033.nt.2	23466	23466.01	Chip 1	2452-2508
DEX0367_033.nt.2	23466	23466.02	Chip 1	2452-2508
DEX0367_034.nt.1	17278	17278.01	Chip 1	991-1050
DEX0367_034.nt.1	17278	17278.02	Chip 1	991-1050
DEX0367_034.nt.2	17278	17278.01	Chip 1	713-772
DEX0367_034.nt.2	17278	17278.02	Chip 1	713-772
DEX0367_035.nt.1	20773	20773.01	Chip 1	2724-2783
DEX0367_035.nt.1	20773	20773.02	Chip 1	2724-2783
DEX0367_035.nt.2	20773	20773.01	Chip 1	2810-2869
DEX0367_035.nt.2	20773	20773.02	Chip 1	2810-2869
DEX0367_035.nt.3	20773	20773.01	Chip 1	1153-1212
DEX0367_035.nt.3	20773	20773.02	Chip 1	1153-1212
DEX0367_036.nt.1	18520	18520.01	Chip 1	4577-4518
DEX0367_036.nt.1	18520	18520.02	Chip 1	4577-4518
DEX0367_036.nt.1	22734	22734.01	Chip 1	4577-4518
DEX0367_036.nt.1	22734	22734.02	Chip 1	4577-4518
DEX0367_037.nt.1	2023	2023	Chip 2	442-501
DEX0367_037.nt.1	5327	5327	Chip 2	442-501
DEX0367_037.nt.1	5328	5328	Chip 2	402-461
DEX0367_038.nt.1	10722	10722.01	Chip 1	2454-2513
DEX0367_038.nt.1	10722	10722.02	Chip 1	2454-2513
DEX0367_039.nt.1	103385	103385.01	Chip 1	926-985
DEX0367_039.nt.1	103385	103385.02	Chip 1	926-985
DEX0367_039.nt.2	103385	103385.01	Chip 1	1155-1214
DEX0367_039.nt.2	103385	103385.02	Chip 1	1155-1214
DEX0367_039.nt.3	103385	103385.01	Chip 1	1034-1093
DEX0367_039.nt.3	103385	103385.02	Chip 1	1034-1093
DEX0367_040.nt.1	92327	92327.01	Chip 1	177-236
DEX0367_040.nt.1	92327	92327.02	Chip 1	177-236
DEX0367_040.nt.2	92327	92327.01	Chip 1	5349-5408
DEX0367_040.nt.2	92327	92327.02	Chip 1	5349-5408
DEX0367_040.nt.3	92327	92327.01	Chip 1	142-201
DEX0367_040.nt.3	92327	92327.02	Chip 1	142-201
DEX0367_040.nt.4	92327	92327.01	Chip 1	58-117
DEX0367_040.nt.4	92327	92327.02	Chip 1	58-117
DEX0367_041.nt.1	11575	11575.01	Chip 1	685-744
DEX0367_041.nt.1	11575	11575.02	Chip 1	685-744

DEX0367_041.nt.1	17486	17486.01	Chip 1	680-739
DEX0367_041.nt.1	17486	17486.02	Chip 1	680-739
DEX0367_041.nt.1	17490	17490.01	Chip 1	687-746
DEX0367_041.nt.1	17490	17490.02	Chip 1	687-746
DEX0367_041.nt.2	11575	11575.01	Chip 1	892-951
DEX0367_041.nt.2	11575	11575.02	Chip 1	892-951
DEX0367_041.nt.2	17486	17486.01	Chip 1	887-946
DEX0367_041.nt.2	17486	17486.02	Chip 1	887-946
DEX0367_041.nt.2	17490	17490.01	Chip 1	894-953
DEX0367_041.nt.2	17490	17490.02	Chip 1	894-953
DEX0367_041.nt.3	11575	11575.01	Chip 1	741-800
DEX0367_041.nt.3	11575	11575.02	Chip 1	741-800
DEX0367_041.nt.3	17486	17486.01	Chip 1	736-795
DEX0367_041.nt.3	17486	17486.02	Chip 1	736-795
DEX0367_041.nt.3	17490	17490.01	Chip 1	743-802
DEX0367_041.nt.3	17490	17490.02	Chip 1	743-802
DEX0367_041.nt.4	11575	11575.01	Chip 1	1457-1516
DEX0367_041.nt.4	11575	11575.02	Chip 1	1457-1516
DEX0367_041.nt.4	17486	17486.01	Chip 1	1452-1511
DEX0367_041.nt.4	17486	17486.02	Chip 1	1452-1511
DEX0367_041.nt.4	17490	17490.01	Chip 1	1459-1518
DEX0367_041.nt.4	17490	17490.02	Chip 1	1459-1518
DEX0367_041.nt.5	11575	11575.01	Chip 1	2397-2456
DEX0367_041.nt.5	11575	11575.02	Chip 1	2397-2456
DEX0367_041.nt.5	17486	17486.01	Chip 1	2392-2451
DEX0367_041.nt.5	17486	17486.02	Chip 1	2392-2451
DEX0367_041.nt.5	17490	17490.01	Chip 1	2399-2458
DEX0367_041.nt.5	17490	17490.02	Chip 1	2399-2458
DEX0367_041.nt.6	11575	11575.01	Chip 1	513-572
DEX0367_041.nt.6	11575	11575.02	Chip 1	513-572
DEX0367_041.nt.6	17486	17486.01	Chip 1	508-567
DEX0367_041.nt.6	17486	17486.02	Chip 1	508-567
DEX0367_041.nt.6	17490	17490.01	Chip 1	515-574
DEX0367_041.nt.6	17490	17490.02	Chip 1	515-574
DEX0367_041.nt.7	11575	11575.01	Chip 1	576-635
DEX0367_041.nt.7	11575	11575.02	Chip 1	576-635
DEX0367_041.nt.7	17486	17486.01	Chip 1	571-630
DEX0367_041.nt.7	17486	17486.02	Chip 1	571-630
DEX0367_041.nt.7	17490	17490.01	Chip 1	578-637
DEX0367_041.nt.7	17490	17490.02	Chip 1	578-637
DEX0367_042.nt.1	23542	23542	Chip 2	5118-5177
DEX0367_042.nt.1	23542	23542.1	Chip 2	5118-5177
DEX0367_042.nt.1	23543	23543	Chip 2	5011-5070
DEX0367_042.nt.1	23543	23543.1	Chip 2	5011-5070
DEX0367_042.nt.2	23542	23542	Chip 2	4550-4609
DEX0367_042.nt.2	23542	23542.1	Chip 2	4550-4609
DEX0367_042.nt.2	23543	23543	Chip 2	4443-4502
DEX0367_042.nt.2	23543	23543.1	Chip 2	4443-4502
DEX0367_042.nt.3	23542	23542	Chip 2	2635-2694
DEX0367_042.nt.3	23542	23542.1	Chip 2	2635-2694
DEX0367_042.nt.3	23543	23543	Chip 2	2528-2587
DEX0367_042.nt.3	23543	23543.1	Chip 2	2528-2587
DEX0367_043.nt.1	21505	21505.01	Chip 1	355-414
DEX0367_043.nt.1	21505	21505.02	Chip 1	355-414
DEX0367_043.nt.2	11527	11527.01	Chip 1	467-526
DEX0367_043.nt.2	11527	11527.02	Chip 1	467-526
DEX0367_044.nt.1	21501	21501.01	Chip 1	772-831

DEX0367_044.nt.1	21501	21501.02	Chip 1	772-831
DEX0367_044.nt.1	21511	21511.01	Chip 1	586-645
DEX0367_044.nt.1	21511	21511.02	Chip 1	586-645
DEX0367_044.nt.2	21501	21501.01	Chip 1	1189-1248
DEX0367_044.nt.2	21501	21501.02	Chip 1	1189-1248
DEX0367_044.nt.2	21511	21511.01	Chip 1	1003-1062
DEX0367_044.nt.2	21511	21511.02	Chip 1	1003-1062
DEX0367_045.nt.1	12155	12155.01	Chip 1	309-368
DEX0367_045.nt.1	12155	12155.02	Chip 1	309-368
DEX0367_045.nt.2	12155	12155.01	Chip 1	305-364
DEX0367_045.nt.2	12155	12155.02	Chip 1	305-364
DEX0367_046.nt.1	10702	10702.01	Chip 1	585-645
DEX0367_046.nt.1	10702	10702.02	Chip 1	585-645
DEX0367_047.nt.1	14656	14656.01	Chip 1	463-522
DEX0367_047.nt.1	14656	14656.02	Chip 1	463-522
DEX0367_047.nt.2	14656	14656.01	Chip 1	1299-1350
DEX0367_047.nt.2	14656	14656.02	Chip 1	1299-1350
DEX0367_047.nt.3	14656	14656.01	Chip 1	590-649
DEX0367_047.nt.3	14656	14656.02	Chip 1	590-649
DEX0367_048.nt.1	9958	9958.01	Chip 1	599-540
DEX0367_048.nt.1	9958	9958.02	Chip 1	599-540
DEX0367_049.nt.1	36013	36013.01	Chip 1	382-441
DEX0367_049.nt.1	36013	36013.02	Chip 1	382-441
DEX0367_050.nt.1	17314	17314.01	Chip 1	614-673
DEX0367_050.nt.1	17314	17314.02	Chip 1	614-673
DEX0367_051.nt.1	22573	22573.01	Chip 1	392-451
DEX0367_051.nt.1	22573	22573.02	Chip 1	392-451
DEX0367_052.nt.1	105767	105767	Chip 2	3314-3373
DEX0367_052.nt.1	105767	105767.1	Chip 2	3314-3373
DEX0367_052.nt.1	105768	105768	Chip 2	3274-3333
DEX0367_052.nt.1	105768	105768.1	Chip 2	3274-3333
DEX0367_052.nt.2	105767	105767	Chip 2	1518-1577
DEX0367_052.nt.2	105767	105767.1	Chip 2	1518-1577
DEX0367_052.nt.2	105768	105768	Chip 2	1478-1537
DEX0367_052.nt.2	105768	105768.1	Chip 2	1478-1537
DEX0367_053.nt.1	1168	1168	Chip 2	180-239
DEX0367_053.nt.2	1168	1168	Chip 2	391-450
DEX0367_054.nt.1	11511	11511.01	Chip 1	2111-2170
DEX0367_054.nt.1	11511	11511.02	Chip 1	2111-2170
DEX0367_055.nt.1	23378	23378	Chip 2	684-743
DEX0367_055.nt.1	23378	23378.01	Chip 1	684-743
DEX0367_055.nt.1	23378	23378.02	Chip 1	684-743
DEX0367_055.nt.1	23378	23378.1	Chip 2	684-743

The following sequences were also up-regulated on a breast microarray. Accordingly, in addition to ovarian cancer, the proteins or the nucleic acid may be used as a breast cancer therapeutic or diagnostic target.

SeqID	Master OligoID	oligoID	Location
DEX0367_008.nt.1	23378	23378	69-15
DEX0367_055.nt.1	23378	23378	684-743

5

The following sequences were also up-regulated on a colon microarray. Accordingly, in addition to ovarian cancer, the proteins or the nucleic acid may be used as a colon cancer therapeutic or diagnostic target.

SeqID	Master OligoID	oligoID	Location
DEX0367_008.nt.1	23378	23378	69-15
DEX0367_008.nt.1	23378	23378.1	69-15
DEX0367_013.nt.1	35316	35316	688-742
DEX0367_013.nt.1	35317	35317	668-727
DEX0367_013.nt.2	35316	35316	418-472
DEX0367_013.nt.2	35317	35317	398-457
DEX0367_014.nt.1	34334	34334	3108-3167
DEX0367_014.nt.1	34335	34335	3022-3081
DEX0367_014.nt.2	34334	34334	2527-2586
DEX0367_014.nt.2	34335	34335	2441-2500
DEX0367_055.nt.1	23378	23378	684-743
DEX0367_055.nt.1	23378	23378.1	684-743

5   The following sequences were also up-regulated on a lung microarray. Accordingly, in addition to ovarian cancer, the proteins or the nucleic acid may be used as a lung cancer therapeutic or diagnostic target.

SeqID	Master OligoID	oligoID	Location
DEX0367_008.nt.1	23378	23378	69-15
DEX0367_055.nt.1	23378	23378	684-743

The following sequences were also up-regulated on a prostate microarray. Accordingly, in

10   addition to ovarian cancer, the proteins or the nucleic acid may be used as a prostate cancer therapeutic or diagnostic target.

SeqID	Master OligoID	oligoID	Location
DEX0367_008.nt.1	23378	23378.01	69-15
DEX0367_008.nt.1	23378	23378.01	69-15
DEX0367_008.nt.1	23378	23378.02	69-15
DEX0367_008.nt.1	23378	23378.02	69-15
DEX0367_042.nt.1	23542	23542.01	5118-5177
DEX0367_042.nt.1	23542	23542.02	5118-5177
DEX0367_042.nt.2	23542	23542.01	4550-4609
DEX0367_042.nt.2	23542	23542.02	4550-4609
DEX0367_042.nt.3	23542	23542.01	2635-2694
DEX0367_042.nt.3	23542	23542.02	2635-2694
DEX0367_055.nt.1	23378	23378.01	684-743
DEX0367_055.nt.1	23378	23378.01	684-743
DEX0367_055.nt.1	23378	23378.02	684-743
DEX0367_055.nt.1	23378	23378.02	684-743

The mapping of the nucleic acid ("NT") SEQ ID NO; DEX ID; chromosomal location (if known); open reading frame (ORF) location; amino acid ("AA") SEQ ID NO and AA

15   DEX ID are shown in the table below:

NT_SEQ No	NT_SEQID	Chromo Map	ORF_Loc	AA_SEQ NO	AA_SEQID
1	DEX0367_001.nt.1	9p13.3	85-965	100	DEX0367_001.aa.1
2	DEX0367_001.nt.2	9p13.3	85-599	101	DEX0367_001.aa.2
3	DEX0367_002.nt.1	Xp11.23	322-1035	102	DEX0367_002.aa.1
4	DEX0367_003.nt.1	1q32.1	111-657	103	DEX0367_003.aa.1
5	DEX0367_004.nt.1	*17p11.2	217-918	104	DEX0367_004.aa.1
6	DEX0367_005.nt.1	10p11.23	130-479	105	DEX0367_005.aa.1
7	DEX0367_006.nt.1	3q29	2974-5074	106	DEX0367_006.aa.1
8	DEX0367_006.nt.2	3q29	2968-5257	107	DEX0367_006.aa.2
9	DEX0367_007.nt.1	2q14.3	853-1270	108	DEX0367_007.aa.1
10	DEX0367_007.nt.2	2q14.3	853-1270	108	DEX0367_007.aa.1
11	DEX0367_008.nt.1	9q21.33	729-1268	109	DEX0367_008.aa.1
12	DEX0367_009.nt.1	19q13.33	1-885	110	DEX0367_009.aa.1
13	DEX0367_010.nt.1	4q27	868-1141	111	DEX0367_010.aa.1
14	DEX0367_011.nt.1	20p12.2	1-1123	112	DEX0367_011.aa.1
15	DEX0367_012.nt.1	*6p21.32	151-370	113	DEX0367_012.aa.1
16	DEX0367_012.nt.2	*6p21.32	151-490	114	DEX0367_012.aa.2
17	DEX0367_013.nt.1	19q13.2	169-657	115	DEX0367_013.aa.1
18	DEX0367_013.nt.2	19q13.2	1-387	116	DEX0367_013.aa.2
19	DEX0367_014.nt.1	1q32.1	206-977	117	DEX0367_014.aa.1
20	DEX0367_014.nt.2	1q32.1	206-1415	118	DEX0367_014.aa.2
21	DEX0367_015.nt.1	12p12.3	10-669	119	DEX0367_015.aa.1
22	DEX0367_015.nt.2	12p12.3	123-642	120	DEX0367_015.aa.2
23	DEX0367_016.nt.1	1q42.2	82-1870	121	DEX0367_016.aa.1
24	DEX0367_016.nt.2	1q42.2	73-1192	122	DEX0367_016.aa.2
25	DEX0367_017.nt.1	12q13.2	1-258	123	DEX0367_017.aa.1
26	DEX0367_018.nt.1	1p31.1	103-871	124	DEX0367_018.aa.1
27	DEX0367_019.nt.1	1p34.1	101-626	125	DEX0367_019.aa.1
28	DEX0367_019.nt.2	1p34.1	381-1158	126	DEX0367_019.aa.2
29	DEX0367_020.nt.1	*	154-553	127	DEX0367_020.aa.1
30	DEX0367_021.nt.1	9q34.11	208-1273	128	DEX0367_021.aa.1
31	DEX0367_021.nt.2	9q34.11	681-2151	129	DEX0367_021.aa.2
32	DEX0367_022.nt.1	11q13.4	65-929	130	DEX0367_022.aa.1
33	DEX0367_023.nt.1	19p13.11	364-796	131	DEX0367_023.aa.1
34	DEX0367_023.nt.2	19p13.11	474-678	132	DEX0367_023.aa.2
35	DEX0367_024.nt.1	1p36.11	1-250	133	DEX0367_024.aa.1
36	DEX0367_024.nt.2	1p36.11	189-1272	134	DEX0367_024.aa.2
37	DEX0367_024.nt.3	1p36.11	189-1590	135	DEX0367_024.aa.3
38	DEX0367_024.nt.4	1p36.11	189-1176	136	DEX0367_024.aa.4
39	DEX0367_025.nt.1	19p13.12	108-645	137	DEX0367_025.aa.1
40	DEX0367_025.nt.2	19p13.12	108-675	138	DEX0367_025.aa.2
41	DEX0367_025.nt.3	19p13.12	91-327	139	DEX0367_025.aa.3
42	DEX0367_026.nt.1	19q13.42	13-1425	140	DEX0367_026.aa.1
43	DEX0367_027.nt.1	7q11.21	608-959	141	DEX0367_027.aa.1
44	DEX0367_028.nt.1	2p13.3	486-1569	142	DEX0367_028.aa.1
45	DEX0367_028.nt.2	2p13.3	468-1002	143	DEX0367_028.aa.2
46	DEX0367_029.nt.1	17q25.1	474-1617	144	DEX0367_029.aa.1
47	DEX0367_029.nt.2	17q25.1	474-2514	145	DEX0367_029.aa.2
48	DEX0367_029.nt.3	17q25.1	474-1617	144	DEX0367_029.aa.1
49	DEX0367_029.nt.4	17q25.1	474-1617	144	DEX0367_029.aa.1
50	DEX0367_030.nt.1	*	1-236	146	DEX0367_030.aa.1
51	DEX0367_031.nt.1	2q24.2	16-128	147	DEX0367_031.aa.1
52	DEX0367_032.nt.1	9q21.11	140-788	148	DEX0367_032.aa.1
53	DEX0367_033.nt.1	9p24.3	2985-5847	149	DEX0367_033.aa.1

54	DEX0367_033.nt.2	9p24.3	2962-5149	150	DEX0367_033.aa.2
55	DEX0367_034.nt.1	16p11.2	196-1134	151	DEX0367_034.aa.1
56	DEX0367_034.nt.2	16p11.2	82-630	152	DEX0367_034.aa.2
57	DEX0367_035.nt.1	12p13.31	248-2156	153	DEX0367_035.aa.1
58	DEX0367_035.nt.2	12p13.31	248-749	154	DEX0367_035.aa.2
59	DEX0367_035.nt.3	12p13.31	1-585	155	DEX0367_035.aa.3
60	DEX0367_036.nt.1	7q22.3	37-764	156	DEX0367_036.aa.1
61	DEX0367_037.nt.1	1p34.1	160-946	157	DEX0367_037.aa.1
62	DEX0367_038.nt.1	15q15.1	196-1696	158	DEX0367_038.aa.1
63	DEX0367_039.nt.1	10q22.1	100-467	159	DEX0367_039.aa.1
64	DEX0367_039.nt.2	10q22.1	330-696	160	DEX0367_039.aa.2
65	DEX0367_039.nt.3	10q22.1	269-575	161	DEX0367_039.aa.3
66	DEX0367_040.nt.1	*	58-389	162	DEX0367_040.aa.1
67	DEX0367_040.nt.2	*	295-4752	163	DEX0367_040.aa.2
68	DEX0367_040.nt.3	*	88-352	164	DEX0367_040.aa.3
69	DEX0367_040.nt.4	*	1-253	165	DEX0367_040.aa.4
70	DEX0367_041.nt.1	9q34.3	112-583	166	DEX0367_041.aa.1
71	DEX0367_041.nt.2	9q34.3	112-790	167	DEX0367_041.aa.2
72	DEX0367_041.nt.3	9q34.3	112-544	168	DEX0367_041.aa.3
73	DEX0367_041.nt.4	9q34.3	112-1354	169	DEX0367_041.aa.4
74	DEX0367_041.nt.5	9q34.3	112-1345	170	DEX0367_041.aa.5
75	DEX0367_041.nt.6	9q34.3	2-410	171	DEX0367_041.aa.6
76	DEX0367_041.nt.7	9q34.3	1-455	172	DEX0367_041.aa.7
77	DEX0367_042.nt.1	20p12.1	298-3564	173	DEX0367_042.aa.1
78	DEX0367_042.nt.2	20p12.1	298-3564	173	DEX0367_042.aa.1
79	DEX0367_042.nt.3	20p12.1	298-1863	174	DEX0367_042.aa.3
80	DEX0367_043.nt.1	*	1-499	175	DEX0367_043.aa.1
81	DEX0367_043.nt.2	*	1-790	176	DEX0367_043.aa.2
82	DEX0367_044.nt.1	19q13.2	352-991	177	DEX0367_044.aa.1
83	DEX0367_044.nt.2	19q13.2	769-1498	178	DEX0367_044.aa.2
84	DEX0367_045.nt.1	20q11.23	7-138	179	DEX0367_045.aa.1
85	DEX0367_045.nt.2	20q11.23	1-107	180	DEX0367_045.aa.2
86	DEX0367_046.nt.1	4q22.1	70-439	181	DEX0367_046.aa.1
87	DEX0367_047.nt.1	1q42.12	1-93	182	DEX0367_047.aa.1
88	DEX0367_047.nt.2	1q42.12	1-93	182	DEX0367_047.aa.1
89	DEX0367_047.nt.3	1q42.12	1-93	182	DEX0367_047.aa.1
90	DEX0367_048.nt.1	*	444-630	183	DEX0367_048.aa.1
91	DEX0367_049.nt.1	16p13.3	1-492	184	DEX0367_049.aa.1
92	DEX0367_050.nt.1	17q21.33	802-1389	185	DEX0367_050.aa.1
93	DEX0367_051.nt.1	*	88-298	186	DEX0367_051.aa.1
94	DEX0367_052.nt.1	8p23.1	136-334	187	DEX0367_052.aa.1
95	DEX0367_052.nt.2	8p23.1	19-102	188	DEX0367_052.aa.2
96	DEX0367_053.nt.1	Xq28	83-548	189	DEX0367_053.aa.1
97	DEX0367_053.nt.2	Xq28	1-817	190	DEX0367_053.aa.2
98	DEX0367_054.nt.1	2p21	1182-1989	191	DEX0367_054.aa.1
99	DEX0367_055.nt.1	7p22.1	1-149	192	DEX0367_055.aa.1

For the polypeptides of the invention, the following attributes were found, epitopes, post translational modifications, signal peptides and transmembrane domains. Antigenicity (Epitope) prediction was performed through the antigenic module in the

5 EMBOSS package. Rice, P., EMBOSS: The European Molecular Biology Open Software Suite, *Trends in Genetics* 16(6): 276-277 (2000). The antigenic module predicts

potentially antigenic regions of a protein sequence, using the method of Kolaskar and
Tongaonkar. Kolaskar, AS and Tongaonkar, PC., A semi-empirical method for prediction
of antigenic determinants on protein antigens, *FEBS Letters* 276: 172-174 (1990). ·
Examples of post-translational modifications (PTMs) and other motifs of the OSPs of this

5 invention are listed below. In addition, antibodies that specifically bind such post-
translational modifications may be useful as a diagnostic or as therapeutic. The PTMs and ·
other motifs were predicted by using the ProSite Dictionary of Proteins Sites and Patterns ·
(Bairoch *et al.*, *Nucleic Acids Res.* 25(1):217-221 (1997)), the following motifs, including ·
PTMs, were predicted for the OSPs of the invention. The signal peptides were detected by

10 using the SignalP 2.0, *see* Nielsen *et al.*, *Protein Engineering* 12, 3-9 (1999). Prediction
of transmembrane helices in proteins was performed by the application TMHMM 2.0, ·
"currently the best performing transmembrane prediction program", according to authors
(Krogh *et al.*, *Journal of Molecular Biology*, 305(3):567-580, (2001); Moller *et al.*,
Bioinformatics, 17(7):646-653, (2001); Sonnhammer, *et al.*, *A hidden Markov model for*

15 *predicting transmembrane helices in protein sequences* in Glasgow, *et al.* Ed. Proceedings ·
of the Sixth International Conference on Intelligent Systems for Molecular Biology, pages ·
175-182, Menlo Park, CA, 1998. AAAI Press. The PSORT II program may also be used
to predict cellular localizations. Horton et al., *Intelligent Systems for Molecular Biology* 5: ·
147-152 (1997). The table below includes the following sequence annotations: Signal

20 peptide presence; TM (number of membrane domain, topology in orientation and
position); Amino acid location and antigenic index (location, AI score, length); PTM and
other motifs (type, amino acid residue locations).

SeqID	Signal P	TMHMM	Antigenicity	PTM
DEX0367_001 .aa.1	y	6 -i26-48o58-80i101-123o157 - 179i192 - 214o242 -261i	167-185,1.29; 192-204,1.272; 16-50,1.213; 242-272,1.187; 98-129,1.153; 281-289,1.149; 58-96,1.139; 5-14,1.102; 132-139,1.098; 222-228,1.072; 145-153,1.065; 211-217,1.055; 230-238,1.041	Asn_Glycosylation 146-149; Ck2_Phospho_Site 280-283; Myristyl 56-61, 59-64, 75-80, 80-85, 116-121, 122-127, 150-155, 208-213, 258-263; Mip 86-94; Prokar_Lipoprotein 35-45, 262-272;
DEX0367_001 .aa.2	y	3 -i26-48o58-80i101-123o	16-50,1.213; 98-129,1.153; 58-96,1.139; 5-14,1.102; 132-139,1.098; 145-	Asn_Glycosylation 146-149; Myristyl 56-61,. 59-64, 75-80, 80-85, 116-121, 122-127, 150-155;

			153,1.065	Mip 86-94; Prokar_Lipoprotein 35-45;
DEX0367_002 .aa.1	N	0 -o	164-217,1.189; 43-61,1.159; 89-114,1.145; 116-138,1.139; 64-74,1.136; 148-161,1.094; 4-11,1.088; 14-30,1.076; 79-85,1.054; 33-39,1.034	Ck2_Phospho_Site 46-49, 210-213; Leucine_Zipper 162-183; Myristyl 9-14, 192-197, 199-204, 203-208; Pkc_Phospho_Site 34-36, 52-54, 93-95;
DEX0367_003 .aa.1	N	1 - i152-174o	153-177,1.258; 35-52,1.22; 4-16,1.181; 102-110,1.149; 18-31,1.072; 75-82,1.067; 135-141,1.04	Amidation 118-121; Asn_Glycosylation 181-184; Ck2_Phospho_Site 151-154; Myristyl 65-70, 160-165; Pkc_Phospho_Site 34-36, 35-37, 87-89, 90-92, 101-103, 102-104, 128-130, 151-153;
DEX0367_004 .aa.1	N	0 -o	174-212,1.218; 9-44,1.202; 46-65,1.141; 67-81,1.137; 88-148,1.133	Asn_Glycosylation 52-55; Ck2_Phospho_Site 164-167; Myristyl 66-71, 88-93, 117-122, 144-149, 186-191; Lectin_Galactoside 170-191;
DEX0367_005 .aa.1	N	0 -o	56-82,1.155; 30-51,1.14; 104-111,1.084; 9-18,1.083	Amidation 7-10, 106-109; Camp_Phospho_Site 26-29; Ck2_Phospho_Site 29-32, 48-51; Myristyl 30-35, 62-67; Pkc_Phospho_Site 7-9, 37-39, 98-100, 114-116;
DEX0367_006 .aa.1	N	2 - o595-617i667 -689o	121-140,1.245; 445-458,1.242; 579-588,1.236; 593-617,1.232; 462-473,1.227; 629-688,1.194; 224-243,1.183; 102-115,1.182; 690-696,1.182; 23-34,1.175; 245-261,1.17; 51-68,1.161; 70-98,1.151; 429-437,1.15; 265-319,1.148; 475-512,1.147;	Ribosomal_S2_1 250-261; Amidation 24-27, 202-205; Asn_Glycosylation 302-305, 476-479, 574-577, 633-636; Camp_Phospho_Site 20-23, 21-24, 592-595; Ck2_Phospho_Site 41-44, 145-148, 445-448, 547-550; Myristyl 98-103, 261-266, 269-274, 535-540, 562-567, 596-601, 630-635;

			202-222,1.139; 142-174,1.138; 515-531,1.129; 333-341,1.129; 375-383,1.117; 398-405,1.098; 533-553,1.095; 345-354,1.089; 186-193,1.076; 363-369,1.038	Pkc_Phospho_Site 24-26, 33-35, 75- 77, 215-217, 307- 309, 425-427, 519- 521, 532-534, 565- 567, 646-648; Tyr_Phospho_Site 154-160;
DEX0367_006 .aa.2	N	9 - o211- 233i283 - 305o325 - 347i368 - 385o544 - 566i579 - 599o614 - 633i680 - 699o704 -721i	577-601,1.285; 657-705,1.254; 367-394,1.253; 555-569,1.252; 61-74,1.242; 323-353,1.237; 195-204,1.236; 497-519,1.232; 209-233,1.232; 78-89,1.227; 402-459,1.205; 245-313,1.194; 604-653,1.189; 713-724,1.168; 472-480,1.153; 45-53,1.15; 91- 128,1.147; 156- 169,1.143; 523- 553,1.135; 131- 153,1.129; 484- 491,1.123; 737- 744,1.1; 19- 28,1.088; 746- 755,1.027	Asn_Glycosylation 92-95, 190-193, 249-252, 317-320, 368-371, 423-426, 498-501; Camp_Phospho_Site 208-211; Ck2_Phospho_Site 61-64, 325-328, 471-474, 511-514, 681-684; Myristyl 18-23, 19-24, 166- 171, 178-183, 212- 217, 246-251, 696- 701, 717-722; Pkc_Phospho_Site 135-137, 148-150, 181-183, 262-264, 430-432, 511-513, 599-601, 625-627, 681-683, 726-728; Tyr_Phospho_Site 760-766; N4_Mtase 614-619;
DEX0367_007 .aa.1	N	0 -o	88-107,1.22; 19-29,1.121; 45-61,1.116; 76-85,1.064; 125-132,1.063	Ck2_Phospho_Site 34-37, 43-46; Myristyl 26-31, 107-112; Pkc_Phospho_Site 92-94, 136-138;
DEX0367_008 .aa.1	N	0 -o	49-61,1.175; 117-135,1.172; 92-99,1.162; 143-155,1.15; 34-45,1.149; 77-85,1.119; 168-176,1.097; 104-110,1.069; 8-14,1.034	Amidation 143-146; Camp_Phospho_Site 67-70, 132-135; Ck2_Phospho_Site 72-75, 79-82, 111- 114, 115-118; Myristyl 41-46, 138-143; Pkc_Phospho_Site 117-119; Tyr_Phospho_Site 93-100;
DEX0367_009 .aa.1	N	0 -o	133-176,1.151; 78-120,1.147; 4-23,1.139; 244-261,1.123; 194-201,1.095; 39-45,1.069; 58-68,1.056	Camp_Phospho_Site 137-140; Ck2_Phospho_Site 29-32, 61-64, 76- 79, 194-197, 208- 211; Myristyl 125- 130, 206-211, 236- 241, 271-276, 283-

				288; Pkc_Phospho_Site 29-31, 136-138, 275-277;
DEX0367_010 .aa.1	N	0 -o	43-78,1.171; 16-37,1.092	Camp_Phospho_Site 86-89; Ck2_Phospho_Site 62-65; Myristyl 41-46; Pkc_Phospho_Site 17-19, 84-86;
DEX0367_011 .aa.1	N	1 - o329- 351i	4-32,1.238; 328-354,1.187; 258-270,1.167; 289-321,1.16; 156-177,1.157; 38-61,1.154; 231-240,1.152; 73-122,1.145; 189-217,1.126; 139-145,1.097; 272-286,1.083; 247-253,1.064	Amidation 56-59; Asn_Glycosylation 151-154, 225-228; Ck2_Phospho_Site 139-142, 200-203; Myristyl 37-42, 39-44, 88-93, 340- 345; Pkc_Phospho_Site 7-9, 18-20, 233- 235, 246-248; Lamp_1 154-168;
DEX0367_012 .aa.1	N	0 -o	4-12,1.148; 22- 28,1.112; 34- 44,1.09; 55- 69,1.084	Ck2_Phospho_Site 42-45; Pkc_Phospho_Site 42-44; Tyr_Phospho_Site 55-63;
DEX0367_012 .aa.2	N	0 -o	4-50,1.243; 60- 66,1.138; 76- 84,1.122; 87- 103,1.09	Ck2_Phospho_Site 111-114; Myristyl 66-71, 71-76, 101- 106;
DEX0367_013 .aa.1	N	0 -o	152-159,1.159; 112-131,1.118; 46-56,1.115; 5- 18,1.111; 27- 36,1.1; 133- 149,1.098; 71- 77,1.08; 61- 67,1.07	Asn_Glycosylation 87-90, 97-100; Camp_Phospho_Site 105-108; Ck2_Phospho_Site 73-76; Myristyl 44-49, 98-103; Pkc_Phospho_Site 47-49, 61-63, 121- 123; Lectin_Galactoside 100-120;
DEX0367_013 .aa.2	N	0 -o	118-125,1.159; 7-33,1.125; 78- 97,1.118; 99- 115,1.098; 37- 43,1.08	Asn_Glycosylation 53-56, 63-66; Camp_Phospho_Site 71-74; Ck2_Phospho_Site 39-42; Myristyl 64-69; Pkc_Phospho_Site 87-89; Lectin_Galactoside 66-86;
DEX0367_014 .aa.1	N	0 -o	100-128,1.153; 38-45,1.131; 26-34,1.129; 69-82,1.102;	Camp_Phospho_Site 242-245; Ck2_Phospho_Site 25-28, 93-96, 98-

			176-189,1.099; 136-143,1.08; 87-93,1.067; 236-245,1.062; 193-199,1.037	101, 136-139, 137- 140, 220-223; Myristyl 58-63, 104-109, 167-172, 187-192, 205-210, 207-212, 254-259; Pkc_Phospho_Site 59-61;
DEX0367_014 .aa.2	N	0	-o	100-128,1.153; 38-45,1.131; 26-34,1.129; 205-218,1.129; 311-318,1.118; 69-82,1.102; 176-189,1.099; 332-346,1.091; 379-385,1.088; 136-143,1.08; 87-93,1.067; 288-294,1.052; 193-199,1.037	Amidation 278-281, 298-301, 382-385; Asn_Glycosylation 359-362, 393-396; Camp_Phospho_Site 287-290, 355-358; Ck2_Phospho_Site 25-28, 93-96, 98- 101, 136-139, 137- 140, 208-211, 218- 221, 229-232, 251- 254; Myristyl 58- 63, 104-109, 167- 172, 187-192, 205- 210, 238-243; Pkc_Phospho_Site 59-61;
DEX0367_015 .aa.1	N	0	-o	50-60,1.202; 93-115,1.187; 129-146,1.173; 183-198,1.162; 4-34,1.161; 154-173,1.144; 38-45,1.105	Asn_Glycosylation 51-54; Ck2_Phospho_Site 42-45, 178-181; Myristyl 35-40, 54-59, 105-110, 152-157; Pkc_Phospho_Site 205-207, 219-221;
DEX0367_015 .aa.2	N	0	-o	46-68,1.187; 82-99,1.173; 136-151,1.162; 4-10,1.154; 107-126,1.144	Ck2_Phospho_Site 131-134; Myristyl 58-63, 105-110; Pkc_Phospho_Site 158-160, 172-174;
DEX0367_016 .aa.1	N	0	-o	4-11,1.247; 521-544,1.221; 298-306,1.209; 447-458,1.181; 556-564,1.171; 117-125,1.165; 49-73,1.155; 188-197,1.149; 463-478,1.145; 383-391,1.126; 318-331,1.124; 75-82,1.121; 103-110,1.113; 21-44,1.113; 259-275,1.108; 566-573,1.103; 428-435,1.103; 362-379,1.1; 284-290,1.096; 209-229,1.083; 132-143,1.081;	Thiol_Protease_Cys 58-69; Asn_Glycosylation 22-25, 148-151, 253-256, 570-573; Camp_Phospho_Site 192-195, 469-472; Ck2_Phospho_Site 24-27, 52-55, 119- 122, 158-161, 205- 208, 226-229, 255- 258, 293-296, 295- 298, 381-384, 421- 424, 427-430, 448- 451, 481-484, 522- 525; Glycosaminoglycan 486-489; Myristyl 50-55, 149-154, 169-174, 194-199, 216-221, 300-305,

				86-98,1.077; 501-508,1.069; 406-412,1.062; 179-185,1.061; 342-348,1.057; 580-586,1.054; 490-496,1.049	519-524; Pkc_Phospho_Site 119-121, 232-234, 257-259, 295-297, 366-368, 397-399, 475-477, 579-581; Tyr_Phospho_Site 177-184; Ef_Hand 484-496, 514-526;
DEX0367_016 .aa.2	N	0	-o	298-321,1.221; 75-83,1.209; 224-235,1.181; 333-341,1.171; 240-255,1.145; 160-168,1.126; 95-108,1.124; 5-21,1.108; 343-350,1.103; 205-212,1.103; 139-156,1.1; 30-39,1.097; 61-67,1.096; 278-285,1.069; 183-189,1.062; 119-125,1.057; 357-363,1.054; 267-273,1.049	Asn_Glycosylation 347-350; Camp_Phospho_Site 246-249; Ck2_Phospho_Site 70-73, 72-75, 158- 161, 198-201, 204- 207, 225-228, 258- 261, 299-302; Glycosaminoglycan 263-266; Myristyl 48-53, 52-57, 77- 82, 296-301; Pkc_Phospho_Site 72-74, 143-145, 174-176, 252-254, 356-358; Ef_Hand 261-273, 291-303;
DEX0367_017 .aa.1	N	0	-o	4-14,1.174; 63- 82,1.15; 35- 46,1.138; 20- 26,1.04	Ck2_Phospho_Site 24-27, 47-50, 86- 89; Pkc_Phospho_Site 40-42, 73-75;
DEX0367_018 .aa.1	N	0	-o	244-252,1.167; 223-241,1.155; 154-171,1.113; 129-138,1.113; 185-202,1.095; 4-12,1.094; 75- 81,1.064; 86- 94,1.048; 145- 151,1.048; 32- 44,1.04	Asn_Glycosylation 87-90; Myristyl 18-23, 21-26, 22- 27, 23-28, 24-29, 25-30, 26-31, 27- 32, 29-34, 31-36, 51-56, 55-60, 157- 162, 203-208, 230- 235; Pkc_Phospho_Site 164-166, 212-214;
DEX0367_019 .aa.1	N	0	-o	4-15,1.221; 138-155,1.209; 58-65,1.171; 112-129,1.149; 82-95,1.136; 34-41,1.125; 98-110,1.124	Asn_Glycosylation 34-37, 47-50, 71- 74, 77-80; Ck2_Phospho_Site 25-28, 49-52, 75- 78; Myristyl 39- 44, 109-114, 110- 115, 117-122; Pkc_Phospho_Site 36-38, 167-169, 171-173;
DEX0367_019 .aa.2	Y	3	-i12- 34o54- 76i89- 111o	81-112,1.241; 115-127,1.221; 13-78,1.204; 170-177,1.171; 194-207,1.136; 248-255,1.131;	Asn_Glycosylation 146-149, 159-162, 183-186, 189-192; Ck2_Phospho_Site 117-120, 137-140, 161-164, 187-190;

			146-153,1.125; 210-222,1.124; 234-245,1.093	Myristyl 34-39, 79-84, 151-156, 221-226, 222-227; Pkc_Phospho_Site 148-150, 236-238; Prokar_Lipoprotein 70-80;
DEX0367_020 .aa.1	N	0 -o	85-97,1.182; 101-113,1.155; 30-44,1.153; 4- 14,1.131; 75- 83,1.111; 20- 26,1.066; 54- 60,1.053	Ck2_Phospho_Site 20-23, 53-56, 120- 123, 127-130, 128- 131; Myristyl 26- 31, 49-54; Pkc_Phospho_Site 19-21, 30-32, 42- 44, 92-94;
DEX0367_021 .aa.1	N	0 -o	117-150,1.291; 76-110,1.213; 33-44,1.198; 289-311,1.168; 232-240,1.163; 207-225,1.149; 182-199,1.136; 243-265,1.125; 319-338,1.116; 49-55,1.112; 276-282,1.104; 152-174,1.101; 67-74,1.098; 57-63,1.087; 344-350,1.079; 22-28,1.058	Ck2_Phospho_Site 16-19, 67-70, 233- 236; Myristyl 23- 28, 24-29, 179-184, 324-329; Pkc_Phospho_Site 51-53, 339-341, 353-355; Tyr_Phospho_Site 68-75; Prokar_Lipoprotein 88-98; Protein_Kinase_St 150-162;
DEX0367_021 .aa.2	N	0 -o	327-343,1.235; 459-483,1.218; 378-390,1.189; 393-422,1.174; 198-206,1.174; 440-457,1.171; 103-122,1.165; 151-169,1.163; 175-195,1.157; 278-307,1.151; 65-83,1.135; 228-252,1.133; 50-59,1.127; 356-363,1.123; 139-147,1.103; 258-276,1.096; 88-94,1.072; 215-226,1.053	Asn_Glycosylation 131-134; Ck2_Phospho_Site 140-143, 169-172, 312-315, 419-422; Leucine_Zipper 55- 76, 220-241, 227- 248; Myristyl 412- 417, 447-452, 457- 462; Pkc_Phospho_Site 50-52, 65-67, 140- 142, 210-212, 211- 213, 231-233, 439- 441, 458-460;
DEX0367_022 .aa.1	Y	0 -o	8-21,1.242; 171-180,1.184; 209-234,1.163; 129-136,1.123; 33-40,1.121; 140-154,1.114; 91-102,1.114; 268-274,1.108; 65-73,1.106; 189-197,1.098; 109-127,1.096;	Amidation 107-110; Asn_Glycosylation 168-171, 208-211, 248-251; Ck2_Phospho_Site 53-56, 223-226; Myristyl 87-92, 219-224, 252-257; Pkc_Phospho_Site 31-33, 170-172, 198-200, 243-245,

			54-60,1.089; 236-244,1.078	253-255;
DEX0367_023 .aa.1	N	O -o	101-120,1.206; 4-18,1.179; 46-67,1.134; 83-93,1.128; 129-140,1.111; 73-79,1.044	Amidation 30-33; Asn_Glycosylation 102-105; Ck2_Phospho_Site 104-107; Myristyl 73-78, 80-85, 84-89, 124-129, 131-136; Pkc_Phospho_Site 76-78, 77-79, 141-143;
DEX0367_023 .aa.2	N	O -o	52-64,1.187; 11-21,1.128; 29-46,1.092	Asn_Glycosylation 30-33; Ck2_Phospho_Site 49-52; Myristyl 12-17;
DEX0367_024 .aa.1	N	O -o	4-11,1.131; 15-23,1.096; 45-51,1.09	Glycosaminoglycan 81-84;
DEX0367_024 .aa.2	N	O -o	239-267,1.186; 13-34,1.181; 209-218,1.167; 309-318,1.162; 179-186,1.153; 73-108,1.131; 290-300,1.12; 121-147,1.118; 4-10,1.11; 193-199,1.098; 169-175,1.09; 112-119,1.083; 46-55,1.071; 341-347,1.071	Amidation 222-225, 330-333; Asn_Glycosylation 113-116; Camp_Phospho_Site 83-86; Ck2_Phospho_Site 23-26, 49-52; Myristyl 195-200, 208-213, 268-273; Pkc_Phospho_Site 53-55, 61-63;
DEX0367_024 .aa.3	N	O -o	337-393,1.213; 239-267,1.186; 399-417,1.182; 13-34,1.181; 209-218,1.167; 309-318,1.162; 179-186,1.153; 422-434,1.145; 73-108,1.131; 290-300,1.12; 121-147,1.118; 4-10,1.11; 193-199,1.098; 169-175,1.09; 112-119,1.083; 46-55,1.071	Amidation 222-225, 330-333; Asn_Glycosylation 113-116; Camp_Phospho_Site 83-86; Ck2_Phospho_Site 23-26, 49-52; Glycosaminoglycan 465-468; Myristyl 195-200, 208-213, 268-273, 364-369; Pkc_Phospho_Site 53-55, 61-63, 351-353, 377-379, 404-406;
DEX0367_024 .aa.4	N	O -o	239-267,1.186; 13-34,1.181; 209-218,1.167; 309-318,1.162; 179-186,1.153; 73-108,1.131; 290-300,1.12; 121-147,1.118; 4-10,1.11; 193-	Amidation 222-225; Asn_Glycosylation 113-116; Camp_Phospho_Site 83-86; Ck2_Phospho_Site 23-26, 49-52; Myristyl 195-200, 208-213, 268-273;

			199,1.098; 169-175,1.09; 112-119,1.083; 46-55,1.071	Pkc_Phospho_Site 53-55, 61-63;
DEX0367_025.aa.1	y	1 -i22-440	19-46,1.26; 140-149,1.16; 54-73,1.132; 5-12,1.131; 81-101,1.08; 161-171,1.056; 151-157,1.051; 123-135,1.041	Amidation 178-181; Asn_Glycosylation 70-73, 97-100; Ck2_Phospho_Site 7-10, 161-164; Pkc_Phospho_Site 50-52, 150-152; Tyr_Phospho_Site 152-158;
DEX0367_025.aa.2	y	1 -i22-440	19-46,1.26; 169-185,1.221; 148-157,1.16; 131-144,1.136; 54-73,1.132; 5-12,1.131; 81-101,1.08; 113-128,1.071; 159-165,1.051	Asn_Glycosylation 70-73, 97-100; Ck2_Phospho_Site 7-10, 169-172; Pkc_Phospho_Site 50-52, 158-160; Tyr_Phospho_Site 160-166;
DEX0367_025.aa.3	y	0 -o	48-65,1.202; 6-14,1.07	Myristyl 11-16, 18-23, 19-24, 25-30, 32-37, 39-44, 44-49, 53-58; Pkc_Phospho_Site 73-75;
DEX0367_026.aa.1	N	0 -o	4-37,1.232; 428-442,1.178; 397-414,1.153; 201-208,1.141; 271-296,1.135; 228-235,1.134; 89-104,1.13; 63-79,1.127; 113-132,1.127; 243-257,1.122; 447-453,1.113; 344-351,1.104; 302-317,1.091; 41-56,1.079; 383-393,1.077; 419-425,1.071; 364-370,1.055; 172-178,1.052; 374-380,1.046	Asn_Glycosylation 110-113, 245-248; Ck2_Phospho_Site 64-67, 93-96, 139-142, 148-151, 161-164, 174-177, 342-345, 428-431, 444-447, 453-456; Myristyl 338-343, 385-390, 403-408, 409-414; Pkc_Phospho_Site 57-59, 120-122, 139-141, 250-252, 263-265, 342-344; Tyr_Phospho_Site 61-68;
DEX0367_027.aa.1	N	0 -o	69-80,1.147; 4-13,1.11; 27-37,1.085; 50-61,1.08; 82-89,1.049	Amidation 44-47; Camp_Phospho_Site 53-56, 59-62; Ck2_Phospho_Site 41-44; Pkc_Phospho_Site 48-50, 52-54, 57-59, 62-64, 102-104, 106-108;
DEX0367_028.aa.1	N	0 -o	348-357,1.265; 274-284,1.206; 189-197,1.193; 61-70,1.189;	Annexin 77-129, 149-201, 233-285, 308-360; Asn_Glycosylation

			232-240,1.187; 78-85,1.168; 147-157,1.162; 245-252,1.15; 127-136,1.145; 309-317,1.126; 208-215,1.12; 117-124,1.107; 324-340,1.091; 289-295,1.073	291-294; Camp_Phospho_Site 168-171, 259-262; Ck2_Phospho_Site 76-79, 78-81, 163-166, 181-184, 262-265, 277-280, 320-323, 349-352; Myristyl 11-16, 34-39, 149-154, 308-313; Pkc_Phospho_Site 12-14, 27-29, 38-40, 39-41, 262-264, 269-271, 303-305;
DEX0367_028 .aa.2	N	0 -o	165-174,1.265; 91-101,1.206; 49-57,1.187; 4-14,1.157; 62-69,1.15; 126-134,1.126; 25-32,1.12; 141-157,1.091; 106-112,1.073	Annexin 50-102, 125-177; Asn_Glycosylation 108-111; Camp_Phospho_Site 76-79; Ck2_Phospho_Site 79-82, 94-97, 137-140, 166-169; Myristyl 125-130; Pkc_Phospho_Site 79-81, 86-88, 120-122;
DEX0367_029 .aa.1	N	0 -o	175-185,1.225; 124-165,1.203; 362-370,1.179; 25-49,1.169; 280-315,1.154; 73-90,1.133; 9-15,1.13; 111-121,1.127; 323-341,1.121; 95-102,1.121; 253-265,1.112; 201-211,1.092; 55-61,1.085; 350-357,1.074; 243-250,1.061; 269-278,1.061	Asn_Glycosylation 56-59, 349-352; Ck2_Phospho_Site 123-126, 174-177, 233-236, 235-238, 313-316, 323-326, 357-360; Myristyl 60-65, 229-234, 362-367; Pkc_Phospho_Site 73-75, 76-78, 301-303, 354-356, 363-365; Tyr_Phospho_Site 303-309;
DEX0367_029 .aa.2	N	0 -o	175-185,1.225; 124-165,1.203; 473-482,1.179; 25-49,1.169; 578-589,1.155; 280-315,1.154; 394-416,1.152; 368-384,1.143; 424-465,1.143; 489-499,1.139; 73-90,1.133; 9-15,1.13; 111-121,1.127; 323-355,1.122; 95-102,1.121; 521-528,1.118; 253-	Asn_Glycosylation 56-59; Ck2_Phospho_Site 123-126, 174-177, 233-236, 235-238, 313-316, 323-326, 545-548, 645-648, 658-661; Leucine_Zipper 458-479; Myristyl 60-65, 229-234, 354-359, 366-371, 512-517; Pkc_Phospho_Site 73-75, 76-78, 301-303, 518-520, 531-

			265,1.112; 539-548,1.107; 601-618,1.106; 201-211,1.092; 623-630,1.089; 55-61,1.085; 655-665,1.082; 386-392,1.069; 557-564,1.066; 243-250,1.061; 269-278,1.061; 507-513,1.056; 633-639,1.052	533, 560-562, 593-595, 636-638, 637-639, 657-659; Tyr_Phospho_Site 303-309;
DEX0367_030 .aa.1	N	0 -i	7-34,1.196; 47-54,1.103	Ck2_Phospho_Site 58-61, 63-66;
DEX0367_031 .aa.1	N	0 -i	5-28,1.145	
DEX0367_032 .aa.1	N	0 -o	104-115,1.259; 19-26,1.168; 68-76,1.148; 192-200,1.111; 58-66,1.107; 39-53,1.106; 85-94,1.106; 130-138,1.1; 142-151,1.093; 157-178,1.089; 202-212,1.068	Atp_Gtp_A 54-61; Ck2_Phospho_Site 12-15, 33-36, 35-38, 87-90, 116-119, 172-175; Glycosaminoglycan 41-44; Myristyl 57-62, 111-116, 149-154; Pkc_Phospho_Site 73-75, 116-118, 215-217;
DEX0367_033 .aa.1	N	5 -i21-43o56-78i189-211o266 – 288i295 -317o	141-213,1.251; 753-782,1.251; 570-587,1.245; 52-80,1.244; 655-680,1.23; 292-319,1.22; 21-46,1.201; 531-544,1.196; 624-652,1.188; 121-132,1.185; 884-898,1.185; 787-805,1.18; 818-881,1.176; 589-601,1.168; 348-380,1.166; 609-619,1.166; 721-739,1.162; 99-116,1.162; 512-524,1.154; 255-286,1.151; 321-340,1.148; 409-462,1.144; 492-507,1.141; 924-935,1.123; 243-252,1.116; 4-19,1.116; 84-92,1.108; 388-401,1.1; 551-568,1.095; 914-920,1.094; 691-697,1.081; 810-	Asn_Glycosylation 56-59, 233-236, 746-749, 752-755, 787-790; Ck2_Phospho_Site 508-511, 551-554, 702-705, 719-722, 884-887, 909-912, 939-942, 941-944; Myristyl 135-140, 245-250, 656-661, 700-705; Pkc_Phospho_Site 3-5, 144-146, 166-168, 235-237, 264-266, 322-324, 337-339, 356-358, 551-553, 627-629, 721-723, 760-762, 789-791, 813-815, 840-842, 884-886, 909-911, 914-916, 917-919, 943-945; Tyr_Phospho_Site 302-308; Prokar_Lipoprotein 664-674;

			816,1.062; 700–708,1.061	
DEX0367_033 .aa.2	N	6 – o130–149i169 – 191o206 – 228i340 – 362o417 – 439i446 –468o	292–364,1.251; 203–243,1.244; 126–152,1.243; 443–470,1.22; 172–197,1.201; 74–83,1.196; 9–27,1.196; 250–283,1.185; 659–673,1.185; 94–120,1.182; 38–69,1.174; 499–531,1.166; 406–437,1.151; 472–491,1.148; 560–613,1.144; 699–710,1.123; 643–656,1.122; 394–403,1.116; 162–170,1.116; 538–552,1.1; 689–695,1.094; 154–160,1.051	Asn_Glycosylation 125–128, 207–210, 384–387; Ck2_Phospho_Site 37–40, 118–121, 659–662, 684–687, 714–717, 716–719; Myristyl 286–291, 396–401; Pkc_Phospho_Site 7–9, 95–97, 158–160, 295–297, 317–319, 386–388, 415–417, 473–475, 488–490, 507–509, 659–661, 684–686, 689–691, 692–694, 718–720; Tyr_Phospho_Site 453–459;
DEX0367_034 .aa.1	N	0 –o	85–105,1.201; 299–309,1.17; 232–243,1.157; 42–55,1.148; 22–40,1.148; 120–175,1.145; 250–278,1.145; 213–225,1.123; 10–20,1.094; 57–71,1.09; 185–198,1.08; 280–289,1.066	Ser_Thr_Phosphatase 116–121; Asn_Glycosylation 255–258; Ck2_Phospho_Site 9–12, 203–206; Myristyl 118–123, 171–176, 272–277; Pkc_Phospho_Site 214–216; Tyr_Phospho_Site 146–154;
DEX0367_034 .aa.2	N	0 –o	47–67,1.252; . 18–36,1.164; 149–160,1.157; 108–118,1.128; 167–179,1.128; 130–142,1.123; 71–81,1.118; 4–12,1.114; 38–44,1.1; 91–97,1.075	Asn_Glycosylation 172–175; Ck2_Phospho_Site 58–61, 121–124; Myristyl 88–93; Pkc_Phospho_Site 100–102, 131–133;
DEX0367_035 .aa.1	y	1 – o426–448i	344–423,1.224; 427–464,1.216; 94–110,1.205; 165–173,1.196; 296–325,1.193; 51–86,1.191; 482–528,1.182; 268–294,1.175; 123–151,1.17; 9–36,1.166; 224–241,1.144; 247–265,1.134; 182–189,1.113; 560–572,1.094;	Tnfr_Ngfr_1 48–85, 88–129; Amidation 387–390; Asn_Glycosylation 45–48, 182–185; Camp_Phospho_Site 124–127; Ck2_Phospho_Site 47–50, 87–90, 92–95, 145–148, 157–160, 194–197, 528–531; Myristyl 28–33, 70–75, 134–139, 168–173, 200–205,

			112-118,1.092; 211-217,1.087; 193-200,1.08; 614-621,1.063; 575-589,1.06; 596-602,1.04	207-212, 248-253, 291-296, 381-386, 401-406, 416-421, 552-557, 557-562, 624-629; Pkc_Phospho_Site 47-49, 74-76, 122-124, 123-125, 187-189, 214-216, 326-328, 336-338, 628-630; Tyr_Phospho_Site 49-55; Prokar_Lipoprotein 441-451;
DEX0367_035.aa.2	Y	0 -o	94-110,1.205; 51-86,1.191; 123-151,1.17; 9-36,1.166; 112-118,1.092	Tnfr_Ngfr_1 48-85, 88-129; Asn_Glycosylation 45-48; Camp_Phospho_Site 124-127; Ck2_Phospho_Site 47-50, 87-90, 92-95, 145-148, 157-160; Myristyl 28-33, 70-75, 134-139; Pkc_Phospho_Site 47-49, 74-76, 122-124, 123-125, 166-168; Tyr_Phospho_Site 49-55;
DEX0367_035.aa.3	y	0 -o	13-25,1.182; 41-87,1.182; 119-131,1.094; 173-180,1.063; 134-148,1.06; 155-161,1.04	Camp_Phospho_Site 14-17; Ck2_Phospho_Site 87-90; Myristyl 111-116, 116-121, 183-188; Pkc_Phospho_Site 187-189;
DEX0367_036.aa.1	N	0 -o	22-32,1.209; 216-238,1.152; 42-50,1.149; 144-168,1.133; 69-78,1.125; 130-141,1.11; 198-204,1.108; 84-96,1.1; 172-179,1.085	Asn_Glycosylation 44-47, 109-112, 147-150; Camp_Phospho_Site 38-41; Ck2_Phospho_Site 64-67, 97-100, 117-120, 174-177, 198-201, 209-212, 211-214; Myristyl 42-47; Pkc_Phospho_Site 117-119, 157-159, 173-175;
DEX0367_037.aa.1	Y	5 -i7-29o49-71i91-113o142-164i176	141-167,1.258; 4-32,1.218; 89-113,1.207; 169-201,1.2; 203-232,1.183; 234-258,1.182; 53-	Ck2_Phospho_Site 32-35, 124-127; Myristyl 15-20, 28-33, 58-63, 60-65, 69-74, 73-78, 109-114, 144-149,

		–198o	74,1.179; 34-40,1.091; 78-84,1.057	146-151, 157-162, 161-166, 164-169, 166-171, 188-193, 195-200; Pkc_Phospho_Site 238-240, 253-255;
DEX0367_038 .aa.1	Y	1 –i12-34o	10-33,1.215; 84-105,1.209; 113-132,1.192; 235-243,1.176; 328-355,1.171; 270-283,1.163; 168-183,1.161; 216-227,1.151; 293-303,1.148; 434-444,1.138; 462-486,1.137; 388-403,1.136; 55-67,1.131; 247-254,1.118; 409-419,1.088; 36-53,1.085; 142-152,1.084; 379-385,1.079; 366-372,1.056; 256-262,1.043	Bpti_Kunitz 283-301, 424-442; Asn_Glycosylation 71-74, 240-243; Ck2_Phospho_Site 79-82, 230-233, 355-358, 379-382; Myristyl 34-39, 48-53, 166-171, 217-222, 286-291, 287-292, 290-295, 315-320, 318-323, 427-432, 428-433, 431-436, 459-464, 462-467, 477-482, 487-492; Pkc_Phospho_Site 85-87, 153-155, 247-249, 258-260, 391-393, 445-447, 455-457, 495-497; Rgd 198-200; Ldlra_1 352-374;
DEX0367_039 .aa.1	N	1 –i93-115o	103-118,1.241; 81-101,1.213; 54-62,1.049	Myristyl 45-50, 59-64;
DEX0367_039 .aa.2	N	1 –i93-115o	103-118,1.241; 81-101,1.213; 54-62,1.049	Myristyl 45-50, 59-64;
DEX0367_039 .aa.3	Y	1 –o72-94i	83-98,1.241; 61-81,1.213; 4-12,1.132; 34-42,1.049	Myristyl 25-30, 39-44;
DEX0367_040 .aa.1	N	1 –o55-77i	53-80,1.248; 4-16,1.142; 29-38,1.133; 90-106,1.111	Asn_Glycosylation 30-33; Ck2_Phospho_Site 106-109; Myristyl 15-20, 72-77, 80-85; Pkc_Phospho_Site 84-86, 85-87; Tyr_Phospho_Site 87-94, 88-94; G_Protein_Receptor 23-39;
DEX0367_040 .aa.2	N	0 –o	904-925,1.198; 759-769,1.186; 282-301,1.181; 1276-1305,1.178; 448-457,1.171; 603-620,1.171; 135-145,1.171; 1061-	Asn_Glycosylation 15-18, 150-153, 171-174, 306-309, 327-330, 343-346, 462-465, 483-486, 618-621, 639-642, 655-658, 774-777, 795-798, 811-814, 930-933, 951-954;

			1081,1.165; 1388-1401,1.161; 673-683,1.16; 1255-1269,1.16; 1405-1417,1.159; 49-61,1.148; 480-511,1.148; 1428-1480,1.146; 1241-1253,1.146; 341-355,1.139; 170-199,1.136; 14-43,1.136; 829-839,1.135; 950-979,1.128; 794-806,1.128; 1105-1114,1.128; 326-335,1.128; 638-647,1.128; 361-374,1.127; 987-998,1.127; 1142-1150,1.127; 517-527,1.127; 649-667,1.12; 1120-1134,1.119; 809-823,1.119; 1161-1180,1.113; 205-221,1.105; 229-257,1.102; 1051-1059,1.101; 715-725,1.093; 1341-1351,1.091; 405-413,1.091; 1195-1210,1.091; 1356-1366,1.086; 114-121,1.085; 537-557,1.079; 853-869,1.079; 873-880,1.078; 93-101,1.078; 1184-1192,1.078; 561-568,1.078; 260-266,1.074; 697-708,1.073; 104-110,1.07; 884-890,1.07; 73-89,1.07; 1311-	Ck2_Phospho_Site 55-58, 87-90, 211-214, 367-370, 445-448, 523-526, 555-558, 679-682, 835-838, 867-870, 991-994; Myristyl 40-45, 52-57, 129-134, 196-201, 208-213, 441-446, 508-513, 664-669, 820-825, 909-914, 959-964, 976-981, 988-993; Pkc_Phospho_Site 11-13, 17-19, 62-64, 107-109, 167-169, 173-175, 218-220, 263-265, 323-325, 329-331, 352-354, 479-481, 530-532, 575-577, 635-637, 641-643, 686-688, 738-740, 751-753, 791-793, 797-799, 842-844, 887-889, 947-949, 953-955; Tyr_Phospho_Site 242-250;

			1324,1.069; 437-443,1.069; 1029-1037,1.068; 271-277,1.063; 126-132,1.062; 1009-1017,1.062; 385-396,1.062; 1040-1046,1.055; 427-433,1.051; 895-901,1.051; 750-756,1.049; 594-600,1.049; 572-578,1.043; 583-589,1.035; 416-422,1.032; 529-535,1.029	
DEX0367_040 .aa.3	N	1 -o33-55i	31-58,1.248; 8-16,1.19; 68-84,1.111	Ck2_Phospho_Site 84-87; Myristyl 50-55, 58-63; Pkc_Phospho_Site 62-64, 63-65; Tyr_Phospho_Site 65-72, 66-72;
DEX0367_040 .aa.4	N	1 -o29-51i	27-54,1.248; 64-80,1.111; 4-13,1.104	Ck2_Phospho_Site 80-83; Myristyl 46-51, 54-59; Pkc_Phospho_Site 58-60, 59-61; Tyr_Phospho_Site 61-68, 62-68;
DEX0367_041 .aa.1	Y	0 -o	8-33,1.189; 60-72,1.178; 115-124,1.178; 126-134,1.13; 94-109,1.076; 45-52,1.052	Asn_Glycosylation 56-59, 83-86; Ck2_Phospho_Site 124-127, 128-131; Myristyl 52-57, 81-86, 105-110, 118-123; Pkc_Phospho_Site 111-113; Lipocalin 38-51;
DEX0367_041 .aa.2	Y	0 -o	8-33,1.189; 60-72,1.178; 150-159,1.178; 161-169,1.13; 210-219,1.123; 199-205,1.08; 94-108,1.076; 137-143,1.07; 116-126,1.06; 45-52,1.052	Asn_Glycosylation 56-59, 83-86; Ck2_Phospho_Site 114-117, 159-162, 163-166; Myristyl 52-57, 81-86, 105-110, 129-134, 134-139, 138-143, 149-154, 153-158; Pkc_Phospho_Site 111-113, 114-116, 194-196, 223-225; Lipocalin 38-51;
DEX0367_041 .aa.3	Y	0 -o	8-33,1.189; 60-72,1.178; 115-124,1.178; 126-135,1.13; 94-	Asn_Glycosylation 56-59, 83-86; Ck2_Phospho_Site 124-127, 128-131,

				109,1.076; 45-52,1.052	138-141; Myristyl 52-57, 81-86, 105-110, 118-123; Pkc_Phospho_Site 111-113; Lipocalin 38-51;
DEX0367_041 .aa.4	Y	0	-o	8-33,1.189; 283-295,1.178; 338-347,1.178; 349-357,1.13; 398-407,1.123; 141-148,1.121; 95-113,1.107; 193-201,1.082; 387-393,1.08; 317-332,1.076; 71-77,1.075; 231-240,1.065; 268-275,1.052; 218-224,1.049; 50-56,1.037; 129-135,1.033	Amidation 91-94, 261-264; Asn_Glycosylation 231-234, 279-282, 306-309; Ck2_Phospho_Site 182-185, 347-350, 351-354; Glycosaminoglycan 250-253; Myristyl 67-72, 84-89, 159-164, 173-178, 175-180, 176-181, 191-196, 195-200, 241-246, 261-266, 275-280, 304-309, 328-333, 341-346; Pkc_Phospho_Site 51-53, 63-65, 73-75, 88-90, 167-169, 221-223, 334-336, 382-384, 411-413; Lipocalin 261-274, 262-274;
DEX0367_041 .aa.5	Y	0	-o	381-396,1.295; 226-238,1.199; 8-33,1.189; 163-172,1.156; 122-135,1.15; 370-376,1.106; 360-368,1.106; 345-351,1.103; 182-188,1.098; 399-407,1.088; 281-288,1.078; 71-77,1.075; 95-112,1.075; 308-317,1.074; 292-298,1.068; 50-56,1.037; 209-215,1.033	Amidation 91-94, 194-197, 290-293, 344-347; Camp_Phospho_Site 196-199; Ck2_Phospho_Site 224-227; Myristyl 67-72, 84-89, 153-158, 158-163, 161-166, 163-168, 191-196, 243-248, 251-256, 253-258, 255-260, 256-261, 269-274, 279-284, 287-292, 304-309, 309-314, 323-328, 327-332, 331-336, 335-340, 341-346; Pkc_Phospho_Site 51-53, 63-65, 73-75, 88-90, 141-143, 147-149, 183-185, 194-196, 211-213;
DEX0367_041 .aa.6	Y	0	-o	46-89,1.223; 8-36,1.181; 120-129,1.123; 109-115,1.08; 38-44,1.052	Asn_Glycosylation 98-101; Myristyl 26-31, 59-64, 76-81; Pkc_Phospho_Site 104-106, 133-135; Prokar_Lipoprotein

				50-60;
DEX0367_041 .aa.7	N	0 -o	75-102,1.154; 11-27,1.137; 110-116,1.09; 38-45,1.081; 47-53,1.067; 137-147,1.065	Asn_Glycosylation 76-79; Ck2_Phospho_Site 51-54; Myristyl 56-61, 67-72, 70-75, 74-79, 79-84, 110-115, 120-125, 143-148; Pkc_Phospho_Site 38-40, 121-123;
DEX0367_042 .aa.1	N	1 -i9-31o	174-186,1.219; 144-172,1.186; 8-31,1.182; 89-127,1.18; 915-930,1.17; 1039-1050,1.161; 1017-1027,1.15; 904-911,1.142; 712-722,1.142; 594-602,1.136; 493-510,1.135; 731-752,1.127; 616-631,1.118; 674-692,1.106; 944-955,1.104; 542-551,1.102; 856-863,1.092; 565-571,1.09; 879-897,1.086; 765-778,1.082; 518-531,1.078; 1006-1012,1.077; 973-986,1.075; 780-793,1.075; 810-816,1.069; 605-611,1.066; 462-468,1.065; 822-829,1.06; 841-849,1.057; 957-964,1.055; 661-667,1.036	Amidation 203-206, 223-226, 233-236, 243-246, 253-256, 263-266, 273-276, 283-286, 293-296, 313-316, 323-326, 333-336, 343-346, 353-356, 373-376, 381-384, 391-394, 401-404, 421-424, 431-434, 441-444, 451-454; Asn_Glycosylation 96-99, 212-215, 936-939; Camp_Phospho_Site 481-484; Ck2_Phospho_Site 37-40, 41-44, 42-45, 140-143, 413-416, 428-431, 566-569, 620-623, 637-640, 669-672, 727-730, 780-783, 786-789, 833-836, 852-855, 1003-1006; Myristyl 191-196, 194-199, 200-205; Pkc_Phospho_Site 37-39, 77-79, 140-142, 169-171, 214-216, 486-488, 530-532, 544-546, 620-622, 727-729, 755-757, 809-811, 907-909;
DEX0367_042 .aa.3	Y	1 -i9-31o	174-186,1.219; 396-413,1.186; 144-172,1.186; 8-31,1.182; 339-377,1.18; 89-127,1.18; 507-515,1.136; 415-423,1.129; 455-464,1.102; 478-484,1.09; 431-444,1.078	Amidation 203-206, 223-226, 233-236, 243-246, 253-256, 263-266, 273-276, 283-286, 293-296, 313-316, 323-326, 333-336; Asn_Glycosylation 96-99, 212-215; Ck2_Phospho_Site 37-40, 41-44, 42-45, 140-143, 390-393, 418-421, 479-

					482; Myristyl 191-196, 194-199, 200-205, 422-427; Pkc_Phospho_Site 37-39, 77-79, 140-142, 169-171, 214-216, 390-392, 443-445, 457-459;
DEX0367_043 .aa.1	Y	0	-o	6-30,1.126; 86-93,1.108; 53-71,1.107; 34-40,1.082; 125-132,1.062; 74-80,1.034	Amidation 157-160, 162-165; Ck2_Phospho_Site 37-40, 71-74, 115-118, 122-125, 124-127, 126-129; Myristyl 46-51, 50-55, 81-86, 83-88; Pkc_Phospho_Site 5-7, 77-79;
DEX0367_043 .aa.2	N	0	-o	174-207,1.173; 139-168,1.129; 6-30,1.126; 118-132,1.11; 86-93,1.108; 210-229,1.107; 53-71,1.107; 103-113,1.095; 34-40,1.082; 74-80,1.034	Amidation 254-257, 259-262; Ck2_Phospho_Site 37-40, 71-74; Myristyl 46-51, 50-55, 81-86, 83-88, 117-122, 178-183; Pkc_Phospho_Site 5-7, 77-79;
DEX0367_044 .aa.1	Y	0	-o	35-47,1.25; 60-70,1.189; 10-28,1.181; 178-188,1.133; 81-91,1.13; 130-136,1.11; 107-113,1.087; 138-145,1.083; 147-153,1.047	Bpti_Kunitz 71-89, 166-184; Asn_Glycosylation 62-65, 99-102; Ck2_Phospho_Site 38-41, 85-88, 95-98; Myristyl 11-16, 29-34, 74-79, 75-80, 78-83, 169-174, 170-175, 173-178, 200-205, 203-208, 208-213, 212-217; Pkc_Phospho_Site 106-108, 177-179, 209-211;
DEX0367_044 .aa.2	Y	0	-o	35-47,1.25; 60-70,1.189; 201-239,1.186; 10-28,1.181; 178-186,1.133; 81-91,1.13; 130-136,1.11; 107-113,1.087; 138-145,1.083; 189-195,1.083; 147-153,1.047	Bpti_Kunitz 71-89, 166-184; Asn_Glycosylation 62-65, 99-102; Ck2_Phospho_Site 38-41, 85-88, 95-98, 219-222; Myristyl 11-16, 29-34, 74-79, 75-80, 78-83, 169-174, 170-175, 173-178; Pkc_Phospho_Site 106-108, 177-179;
DEX0367_045 .aa.1	N	0	-o	5-20,1.129	Asn_Glycosylation 37-40; Myristyl

				11-16, 33-38;
DEX0367_045 .aa.2	Y	0 -i	6-18,1.141; 25-31,1.103	
DEX0367_046 .aa.1	N	0 -o	91-98,1.127; 9-20,1.12; 105-118,1.059; 65-72,1.044; 50-56,1.012	Camp_Phospho_Site 53-56; Ck2_Phospho_Site 28-31, 93-96, 104-107; Myristyl 13-18;
DEX0367_047 .aa.1	N	0 -o	11-26,1.155	Ck2_Phospho_Site 29-32;
DEX0367_048 .aa.1	N	1 -i32-54o	29-58,1.199; 14-24,1.177	Myristyl 32-37;
DEX0367_049 .aa.1	N	0 -o	51-66,1.216; 70-111,1.15; 25-46,1.115; 121-128,1.095; 135-144,1.081	Amidation 70-73; Camp_Phospho_Site 115-118, 151-154; Ck2_Phospho_Site 154-157; Myristyl 15-20, 86-91, 88-93; Pkc_Phospho_Site 9-11, 98-100, 134-136; Prokar_Lipoprotein 86-96;
DEX0367_050 .aa.1	N	0 -o	160-184,1.151; 4-11,1.144; 73-85,1.118; 93-99,1.11; 15-25,1.11; 51-59,1.107; 146-153,1.088; 32-40,1.076	Camp_Phospho_Site 125-128; Ck2_Phospho_Site 161-164; Myristyl 73-78, 76-81, 91-96, 95-100, 109-114, 154-159, 158-163, 162-167, 166-171; Pkc_Phospho_Site 68-70, 124-126; Prokar_Lipoprotein 71-81;
DEX0367_051 .aa.1	N	0 -o	22-34,1.182; 38-50,1.155; 4-18,1.086	Ck2_Phospho_Site 57-60, 64-67, 65-68; Pkc_Phospho_Site 29-31;
DEX0367_052 .aa.1	Y	0 -o	4-30,1.142; 32-40,1.122; 55-62,1.119; 42-48,1.073	Asn_Glycosylation 37-40, 44-47; Ck2_Phospho_Site 28-31; Myristyl 24-29;
DEX0367_052 .aa.2	N	0 -o		
DEX0367_053 .aa.1	Y	0 -o	4-60,1.197; 126-151,1.156; 83-97,1.104; 63-78,1.095; 111-117,1.077; 99-107,1.051	Ck2_Phospho_Site 42-45; Myristyl 85-90; Pkc_Phospho_Site 61-63; Tyr_Phospho_Site 84-91, 110-117;
DEX0367_053 .aa.2	N	0 -o	4-21,1.232; 246-268,1.203;	Asn_Glycosylation 100-103;

			66-90,1.192; 39-50,1.182; 224-236,1.156; 98-158,1.148; 181-195,1.104; 161-176,1.095; 209-215,1.077; 197-205,1.051	Ck2_Phospho_Site 93-96, 124-127, 140-143; Myristyl 43-48, 183-188; Pkc_Phospho_Site 26-28, 39-41, 101- 103, 159-161, 269- 271; Rgd 31-33; Tyr_Phospho_Site 182-189, 208-215;
DEX0367_054 .aa.1	N	1 - o220- 242i	219-244,1.317; 4-24,1.212; 145-153,1.156; 199-209,1.147; 163-172,1.122; 69-76,1.121; 129-139,1.104; 95-105,1.087; 86-93,1.085; 176-182,1.056	Thyroglobulin_1 54-82; Amidation 37-40; Asn_Glycosylation 33-36, 70-73, 157- 160; Ck2_Phospho_Site 86-89, 93-96, 162- 165; Myristyl 14- 19, 34-39, 45-50, 62-67, 71-76, 82- 87, 222-227; Pkc_Phospho_Site 86-88, 95-97, 120- 122, 130-132, 159- 161, 187-189, 248- 250; Tyr_Phospho_Site 110-116;
DEX0367_055 .aa.1	N	0 -o	9-25,1.107	Ck2_Phospho_Site 36-39; Myristyl 21-26; Pkc_Phospho_Site 7-9; Actins_2 34- 42;

Example 2:Relative Quantitation of Gene Expression

Real-Time quantitative PCR with fluorescent Taqman[®] probes is a quantitation detection system utilizing the 5'- 3' nuclease activity of Taq DNA polymerase. The

5　method uses an internal fluorescent oligonucleotide probe (Taqman[®]) labeled with a 5' reporter dye and a downstream, 3' quencher dye. During PCR, the 5'-3' nuclease activity of Taq DNA polymerase releases the reporter, whose fluorescence can then be detected by the laser detector of the Model 7700 Sequence Detection System (PE Applied Biosystems, Foster City, CA, USA). Amplification of an endogenous control is used to standardize the

10　amount of sample RNA added to the reaction and normalize for Reverse Transcriptase (RT) efficiency. Either cyclophilin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), ATPase, or 18S ribosomal RNA (rRNA) is used as this endogenous control. To calculate relative quantitation between all the samples studied, the target RNA levels for one sample were used as the basis for comparative results (calibrator). Quantitation

relative to the "calibrator" can be obtained using the comparative method (User Bulletin #2: ABI PRISM 7700 Sequence Detection System).

The tissue distribution and the level of the target gene are evaluated for every sample in normal and cancer tissues. Total RNA is extracted from normal tissues, cancer

5   tissues, and from cancers and the corresponding matched adjacent tissues. Subsequently, first strand cDNA is prepared with reverse transcriptase and the polymerase chain reaction is done using primers and Taqman® probes specific to each target gene. The results are analyzed using the ABI PRISM 7700 Sequence Detector. The absolute numbers are relative levels of expression of the target gene in a particular tissue compared to the

10   calibrator tissue.

One of ordinary skill can design appropriate primers. The relative levels of expression of the OSNA versus normal tissues and other cancer tissues can then be determined. All the values are compared to the calibrator. Normal RNA samples are commercially available pools, originated by pooling samples of a particular tissue from

15   different individuals.

The relative levels of expression of the OSNA in pairs of matched samples may also be determined. A matched pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual. All the values are compared to the calibrator.

20   In the analysis of matching samples, the OSNAs that show a high degree of tissue specificity for the tissue of interest. These results confirm the tissue specificity results obtained with normal pooled samples. Further, the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual are compared. This comparison provides an indication of specificity for the cancer state (*e.g.* higher

25   levels of mRNA expression in the cancer sample compared to the normal adjacent).

Altogether, the high level of tissue specificity, plus the mRNA overexpression in matched samples tested are indicative of SEQ ID NO: 1-99 being a diagnostic marker for cancer.

Example 3: Protein Expression

30   The OSNA is amplified by polymerase chain reaction (PCR) and the amplified DNA fragment encoding the OSNA is subcloned in pET-21d for expression in E. coli. In addition to the OSNA coding sequence, codons for two amino acids, Met-Ala, flanking the

NH$_2$-terminus of the coding sequence of OSNA, and six histidines, flanking the COOH-terminus of the coding sequence of OSNA, are incorporated to serve as initiating Met/restriction site and purification tag, respectively.

5   An over-expressed protein band of the appropriate molecular weight may be observed on a Coomassie blue stained polyacrylamide gel. This protein band is confirmed by Western blot analysis using monoclonal antibody against 6X Histidine tag.

Large-scale purification of OSP is achieved using cell paste generated from 6-liter bacterial cultures, and purified using immobilized metal affinity chromatography (IMAC). Soluble fractions that are separated from total cell lysate were incubated with a nickel
10   chelating resin. The column is packed and washed with five column volumes of wash buffer. OSP is eluted stepwise with various concentration imidazole buffers.

Example 4: Fusion Proteins

The human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5'and 3' ends of the sequence described below. These primers also should
15   have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector. For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the
20   present invention, isolated by the PCR protocol described in Example 2, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced. If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a
25   heterologous signal sequence. *See, e.g.*, WO 96/34891.

Example 5: Production of an Antibody from a Polypeptide

In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to
30   culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100, μg/ml of streptomycin. The

splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any

suitable myeloma cell line may be employed in accordance with the present invention;

however, it is preferable to employ the parent myeloma cell line (SP20), available from

the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT

5   medium, and then cloned by limiting dilution as described by Wands *et al.*,

Gastroenterology 80: 225-232 (1981).

The hybridoma cells obtained through such a selection are then assayed to identify

clones which secrete antibodies capable of binding the polypeptide. Alternatively,

additional antibodies capable of binding to the polypeptide can be produced in a two-step

10   procedure using anti-idiotypic antibodies. Such a method makes use of the fact that

antibodies are themselves antigens, and therefore, it is possible to obtain an antibody

which binds to a second antibody. In accordance with this method, protein specific

antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such

an animal are then used to produce hybridoma cells, and the hybridoma cells are screened

15   to identify clones which produce an antibody whose ability to bind to the protein-specific

antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic

antibodies to the protein specific antibody and can be used to immunize an animal to

induce formation of further protein-specific antibodies.

20   **Example 6: Method of Determining Alterations in a Gene Corresponding to a**
       Polynucleotide

RNA is isolated from individual patients or from a family of individuals that have

a phenotype of interest. cDNA is then generated from these RNA samples using protocols

known in the art. *See*, Sambrook (2001), *supra*. The cDNA is then used as a template for

25   PCR, employing primers surrounding regions of interest in SEQ ID NO: 1-99. Suggested

PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C;

and 60-120 seconds at 70°C, using buffer solutions described in Sidransky *et al.*, *Science*

252(5006): 706-9 (1991). *See also* Sidransky *et al.*, *Science* 278(5340): 1054-9 (1997).

PCR products are then sequenced using primers labeled at their 5' end with T4

30   polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The

intron-exon borders of selected exons are also determined and genomic PCR products

analyzed to confirm the results. PCR products harboring suspected mutations are then

cloned and sequenced to validate the results of the direct sequencing. PCR products is

cloned into T-tailed vectors as described in Holton *et al., Nucleic Acids Res.*, 19: 1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements may also be determined. Genomic clones are nick-translated with digoxigenin deoxyuridine 5' triphosphate (Boehringer Manheim), and FISH is performed as described in Johnson *et al., Methods Cell Biol.* 35: 73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C-and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. Johnson (1991). Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 7: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

Antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described above. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced. The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbound polypeptide. Next, 50 :1 of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbound conjugate. 75 :1 of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution are added to each well and incubated 1 hour at room temperature.

The reaction is measured by a microtiter plate reader. A standard curve is prepared, using serial dilutions of a control sample, and polypeptide concentrations are plotted on the X-axis (log scale) and fluorescence or absorbance on the Y-axis (linear scale). The concentration of the polypeptide in the sample is calculated using the standard

5 curve.

Example 8: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide

10 alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1, μg/kg/day

15 to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 μg/kg/hour to about 50 mg/kg/hour, either by 1-4 injections per day or by continuous

20 subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are

25 administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which

30 include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semipermeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U. S. Pat. No.3,773,919, EP 58,481, the contents of which are hereby incorporated by reference herein in their entirety), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22: 547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15: 167-277 (1981), and R. Langer, Chem. Tech. 12: 98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE Epstein et al., Proc. Natl. Acad. Sci. USA 82: 3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77: 4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324, the contents of which are hereby incorporated by reference herein in their entirety. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation.

For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides. Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably, the carrier is a parenteral carrier, more preferably, a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic

5 acid; low molecular weight (less than about ten residues) polypeptides, e. g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such

10 as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers

15 will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a

20 stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1 % (w/v) aqueous polypeptide solution, and the resulting mixture

25 is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container (s) can be a notice in the form prescribed

30 by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 9:   Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided above.

Example 10: Method of Treating Increased Levels of the Polypeptide

Antisense or RNAi technology are used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided above.

Example 11: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of

the flask and fresh media (e. g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The

5　monolayer is trypsinized and scaled into larger flasks. pMV-7 (Kirschmeier, P. T. et al., DNA, 7: 219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

10　The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5'and 3'end sequences respectively as set forth in Example 3. Preferably, the 5'primer contains an EcoRI site and the 3'primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA

15　ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB 101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+aml2 packaging cells are grown in tissue culture to

20　confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

25　Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the

30　producer cells. This media is removed and replaced with fresh media.

If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector

that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

5

Example 12: Method of Treatment Using Gene Therapy-In Vivo

Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences

10   into an animal to increase or decrease the expression of the polypeptide.

The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, Tabata H. *et al. Cardiovasc. Res.* 35 (3): 470-479 (1997); Chao J

15 · *et al.* Pharmacol. Res. 35 (6): 517-522 (1997); Wolff J. A. *Neuromuscul. Disord.* 7 (5): 314-318 (1997), Schwartz B. *et al. Gene Ther.* 3 (5): 405-411 (1996); and Tsurumi Y. *et al. Circulation* 94 (12): 3281-3290 (1996); WO 90/11092, WO 98/11779; U. S. Patent No. 5,693,622; 5,705,151; 5,580,859, the contents of which are hereby incorporated by reference herein in their entirety.

20   The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, ovarian, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free

25   from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Felgner P. L. *et al. Ann. NY Acad. Sci.* 772: 126-139 (1995) and Abdallah B. *et al. Biol. Cell* 85 (1): 1-7 (1995)) which

30   can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art

can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the

5   desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, ovarian, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue.

10   Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to

15   the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin

20   fibroblasts. In vivo muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 µg/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and

25   more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of

30   injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to ovarians or bronchial tissues, throat or mucous membranes of the nose. In addition, naked

polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle in vivo is determined as follows. Suitable template DNA for production of mRNA coding for
5   polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by
10   intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin
15   is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different
20   mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice.

The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.
25

Example 13: Transgenic Animals

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e. g., baboons,
30   monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (I. e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., *Appl. Microbiol. Biotechnol.* 40: 691-698 (1994); Carver et al.,

5   *Biotechnology* 11: 1263-1270 (1993); Wright et al., *Biotechnology* 9: 830-834 (1991); and U. S. Pat. No. 4,873,191, the contents of which is hereby incorporated by reference herein in its entirety); retrovirus mediated gene transfer into germ lines (Van der Putten et al., *Proc. Natl. Acad. Sci., USA* 82: 6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., *Cell* 56: 313-321 (1989)); electroporation of

10   cells or embryos (Lo, 1983, *Mol Cell. Biol.* 3: 1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e. g., Ulmer et al., *Science* 259: 1745 (1993); introducing nucleic acid constructs into embryonic pleuripotent stem cells and transferring the stem cells back into the blastocyst; and sperm mediated gene transfer (Lavitrano et al., *Cell* 57: 717-723 (1989). For a review of such techniques, see Gordon,

15   "Transgenic Animals," *Intl. Rev. Cytol.* 115: 171-229 (1989).

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., *Nature* 380: 64-66 (1996); Wilmut et al., *Nature* 385: 810813 (1997)).

20   The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, I. e., mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a

25   particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., *Proc. Natl. Acad. Sci. USA* 89: 6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting

30   is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene

may also be selectively introduced into a particular cell type, thus inactivating the

endogenous gene in only that cell type, by following, for example, the teaching of Gu et

al. (Gu et al., *Science* 265: 103-106 (1994)). The regulatory sequences required for such a

cell-type specific inactivation will depend upon the particular cell type of interest, and will

5   be apparent to those of skill in the art. ·

Once transgenic animals have been generated, the expression of the recombinant

gene may be assayed utilizing standard techniques. Initial screening may be accomplished

by Southern blot analysis or PCR techniques to analyze animal tissues to verify that

integration of the transgene has taken place. The level of mRNA expression of the

10   transgene in the tissues of the transgenic animals may also be assessed using techniques

which include, but are not limited to, Northern blot analysis of tissue samples obtained

from the animal, in situ hybridization analysis, and reverse transcriptase-PCR (rt-PCR).

Samples of transgenic gene-expressing tissue may also be evaluated

immunocytochemically or immunohistochemically using antibodies specific for the

15   transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or

crossbred to produce colonies of the particular animal. Examples of such breeding

strategies include, but are not limited to: outbreeding of founder animals with more than

one integration site in order to establish separate lines; inbreeding of separate lines in

20   order to produce compound transgenics that express the transgene at higher levels because

of the effects of additive expression of each transgene; crossing of heterozygous

transgenic animals to produce animals homozygous for a given integration site in order to

both augment expression and eliminate the need for screening of animals by DNA

analysis; crossing of separate homozygous lines to produce compound heterozygous or

25   homozygous lines; and breeding to place the transgene on a distinct background that is

appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited

to, animal model systems useful in elaborating the biological function of polypeptides of

the present invention, studying conditions and/or disorders associated with aberrant

30   expression, and in screening for compounds effective in ameliorating such conditions

and/or disorders.

Example 14: Knock-Out Animals

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (E. g., see Smithies et al., Nature 317: 230-234 (1985); Thomas & Capecchi, Cell 51: 503512 (1987);

5   Thompson et al., Cell 5: 313-321 (1989)) Alternatively, RNAi technology may be used. For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that

10   express polypeptides of the invention in vivo. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used

15   to generate animal offspring with an inactive targeted gene (e. g., see Thomas & Capecchi 1987 and Thompson 1989, supra). However, this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors that will be apparent to those of skill in the art.

20   In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e. g., knockouts) are administered to a patient in vivo. Such cells may be obtained from the patient (i.e., animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts,

25   bone marrow cells, blood cells (e. g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered in vitro using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral

30   vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be

5   introduced into the patient systemically, e. g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e. g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U. S. Patent No. 5,399,349; and

10   Mulligan & Wilson, U. S. Patent No. 5,460,959, the contents of which are hereby incorporated by reference herein in their entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the

15   cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of

20   polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

While preferred illustrative embodiments of the present invention are described, one skilled in the art will appreciate that the present invention can be practiced by other

25   than the described embodiments, which are presented for purposes of illustration only and not by way of limitation. The present invention is limited only by the claims that follow.

CLAIMS

We claim:

1. An isolated nucleic acid molecule comprising:

   (a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 100-192;

   (b) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-99;

   (c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b); or

   (d) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (a) or (b).

2. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is a cDNA.

3. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is genomic DNA.

4. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is an RNA.

5. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is a mammalian nucleic acid molecule.

6. The nucleic acid molecule according to claim 5, wherein the nucleic acid molecule is a human nucleic acid molecule.

7. A method for determining the presence of a ovarian specific nucleic acid (OSNA) in a sample, comprising the steps of:

   (a) contacting the sample with the nucleic acid molecule of SEQ ID NO: 1-99 under conditions in which the nucleic acid molecule will selectively hybridize to a ovarian specific nucleic acid; and

(b) detecting hybridization of the nucleic acid molecule to a OSNA in the sample, wherein the detection of the hybridization indicates the presence of a OSNA in the sample.

5 8. A vector comprising the nucleic acid molecule of claim 1.

9. A host cell comprising the vector according to claim 8.

10. A method for producing a polypeptide encoded by the nucleic acid molecule

10 according to claim 1, comprising the steps of:

(a) providing a host cell comprising the nucleic acid molecule operably linked to one or more expression control sequences, and

(b) incubating the host cell under conditions in which the polypeptide is produced.

15

11. A polypeptide encoded by the nucleic acid molecule according to claim 1.

12. An isolated polypeptide selected from the group consisting of:

(a) a polypeptide comprising an amino acid sequence with at least 95%

20 sequence identity to of SEQ ID NO: 100-192 ; or

(b) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-99.

25 13. An antibody or fragment thereof that specifically binds to:

(a) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 100-192 ; or

(b) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule

30 comprising a nucleic acid sequence of SEQ ID NO: 1-99.

14. A method for determining the presence of a ovarian specific protein in a sample, comprising the steps of:

(a)   contacting the sample with a suitable reagent under conditions in which the reagent will selectively interact with the ovarian specific protein comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 100-192; and

5   (b)   detecting the interaction of the reagent with a ovarian specific protein in the sample, wherein the detection of binding indicates the presence of a ovarian specific protein in the sample. ,

15.   A method for diagnosing or monitoring the presence and metastases of ovarian

10   cancer in a patient, comprising the steps of:

(a)   determining an amount of:

(i)   a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 100-192; ·

(ii)   a nucleic acid molecule comprising a nucleic acid sequence of SEQ

15   ID NO: 1-99;

(iii)   a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (i) or (ii);

(iv)   a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (i) or (ii);

20   (v)   a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 100-192 ; or

(vi)   a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-99

25   and;

(b)   comparing the amount of the determined nucleic acid molecule or the polypeptide in the sample of the patient to the amount of the ovarian specific marker in a normal control; wherein a difference in the amount of the nucleic acid molecule or the polypeptide in the sample compared to the amount of the nucleic

30   acid molecule or the polypeptide in the normal control is associated with the presence of ovarian cancer.

16. A kit for detecting a risk of cancer or presence of cancer in a patient, said kit comprising a means for determining the presence of:

(a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 100-192;

5 (b) ＞ a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-99;

(c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b); or

(d) a nucleic acid molecule having at least 95% sequence identity to the nucleic 10 acid molecule of (a) or (b); or

(e) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 100-192 ; or

(f) · · a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule 15 comprising a nucleic acid sequence of SEQ ID NO: 1-99.

17. A method of treating a patient with ovarian cancer, comprising the step of administering a composition consisting of:

(a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an 20 amino acid sequence of SEQ ID NO: 100-192;

(b) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-99;

(c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b);

25 (d) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (a) or (b);

(e) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 100-192 ; or

(f) a polypeptide comprising an amino acid sequence encoded by a nucleic 30 acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-99;

to a patient in need thereof, wherein said administration induces an immune response against the ovarian cancer cell expressing the nucleic acid molecule or polypeptide.

18.   A vaccine comprising the polypeptide or the nucleic acid encoding the polypeptide
of claim 12.

5

ABSTRACT OF THE INVENTION

The present invention relates to newly identified nucleic acid molecules and polypeptides present in normal and neoplastic ovarian cells, including fragments, variants and derivatives of the nucleic acids and polypeptides. The present invention also relates to

5   antibodies to the polypeptides of the invention, as well as agonists and antagonists of the polypeptides of the invention. The invention also relates to compositions containing the nucleic acid molecules, polypeptides, antibodies, agonists and antagonists of the invention and methods for the use of these compositions. These uses include identifying, diagnosing, monitoring, staging, imaging and treating ovarian cancer and non-cancerous

10   disease states in ovarian, identifying ovarian tissue, monitoring and identifying and/or designing agonists and antagonists of polypeptides of the invention. The uses also include gene therapy, production of transgenic animals and cells, and production of engineered ovarian tissue for treatment and research.

SEQUENCES

>DEX0367_001.nt.1
AGCCACCAGCGCTGGAGGCCGCTGCTCGCTGCGCCACCGCCTCCCGCCACCCCTGCCCGCCCGACAGCGC
CGCCGCCTGCCCCGCCATGGGTCGACAGAAGGAGCTGGTGTCCCGCTGCGGGGAGATGCTCCACATCCGC
TACCGGCTGCTCCGACAGGCGCTGGCCGAGTGCCTGGGGACCCTCATCCTCGTGATGTTTGGCTGTGGCT
CCGTGGCCCAGGTTGTGCTCAGCCGGGGCACCCACGGTGGTTTCCTCACCATCAACCTGGCCTTTGGCTT
TGCTGTCACTCTGGGCATCCTCATCGCTGGCCAGGTCTCTGGGGCCCACCTGAACCCTGCCGTGACCTTT
GCCATGTGCTTCCTGGCTCGTGAGCCCTGGATCAAGCTGCCCATCTACACCCTGGCACAGACGCTGGGAG
CCTTCTTGGGTGCTGGAATAGTTTTTGGGCTGTATTATGATGCAATCTGGCACTTCGCCGACAACCAGCT
TTTTGTTTCGGGCCCCAATGGCACAGCCGGCATCTTTGCTACCTACCCCTCTGGACACTTGGATATGATC
AATGGCTTCTTTGACCAGTTCATAGGCACAGCCTCCCTTATCGTGTGTGTGCTGGCCATTGTTGACCCCT
ACAACAACCCCGTCCCCCGAGGCCTGGAGGCCTTCACCGTGGGCCTGGTGGTCCTGGTCATTGGCACCTC
CATGGGCTTCAACTCCGGCTATGCCGTCAACCCTGCCCGGGACTTTGGCCCCCGCCTTTTTACAGCCCTT
GCGGGCTGGGGCTCTGCAGTCTTCACGACCGGCCAGCATTGGTGGTGGGTGCCCATCGTGTCCCCACTCC
TGGGCTCCATTGCGGGTGTCTTCGTGTACCAGCTGATGATCGGCTGCCACCTGGAGCAGCCCCCACCCTC
CAACGAGGAAGAGAATGTGAAGCTGGCCCATGTGAAGCACAAGGAGCAGATCTGAGTGGGCAGGGGCCAT
CTCCCCACTCCGCTGCCCTGGCCTTGAGCATCCACTGACTGTCCAAGGGCCACTCCCAAGAAGCCCCCTT
CACGATCCACCCTTTCAGGCTAAGGAGCTCCCTATCTACCCTCACCCCACGAGACAGCCCCTTCAGGATT
TCCACTGGACCTTGCCCAAATAGCACCTTAGGCCACTGCCCCTAAGCTGGGGTGGAACCGGAATTTGGGT
CAATACATCCTTTTGTCTCCCAAGGGAAGAGAATGGGCAGCAGGTATGTGTGTGTGTGCATGTGTGTGCA
TGTGTGTGCATGTGTGTGCAGGGGTGTGTGTGTGTGGGGGGGGTTCCCAGGGCTGGGAGGAGGGGGCCAC
AGCTTAGGTTTGGAGCTCTGGATGTACATACATAAGTAGGAGCAGTGGGACGTGTTTCTGTCATAATGCA
GGCATGAAGGGTGGAGTGAAGTCAGGTCATAAGTTTCATGTTTGCTTTTGTTTTGTTTTGTTTTTAATGT
ATGTAGCAGATGTTACAGTCTTAGGGATCCGGGATGGGAGACCCCACTTTAGAAAGGGTCGTCACTCCTT
TAATCCTCTACTCAACAATGTACTCTTTTACTTTTATATTAAAAAAAAATAAAATAAATATGTGCCTAAAA

>DEX0367_001.nt.2
AGCCACCAGCGCTGGAGGCCGCTGCTCGCTGCGCCACCGCCTCCCGCCACCCCTGCCCGCCCGACAGCGC
CGCCGCCTGCCCCGCCATGGGTCGACAGAAGGAGCTGGTGTCCCGCTGCGGGGAGATGCTCCACATCCGC
TACCGGCTGCTCCGACAGGCGCTGGCCGAGTGCCTGGGGACCCTCATCCTCGTGATGTTTGGCTGTGGCT
CCGTGGCCCAGGTTGTGCTCAGCCGGGGCACCCACGGTGGTTTCCTCACCATCAACCTGGCCTTTGGCTT
TGCTGTCACTCTGGGCATCCTCATCGCTGGCCAGGTCTCTGGGGCCCACCTGAACCCTGCCGTGACCTTT
GCCATGTGCTTCCTGGCTCGTGAGCCCTGGATCAAGCTGCCCATCTACACCCTGGCACAGACGCTGGGAG
CCTTCTTGGGTGCTGGAATAGTTTTTGGGCTGTATTATGATGCAATCTGGCACTTCGCCGACAACCAGCT
TTTTGTTTCGGGCCCCAATGGCACAGCCGGCATCTTTGCTACCTACCCCTCTGGACACTTGGATATGATC
AATGGCTTCTTTGACCAGGTATGGGCTGGGGACGTGTGAGGGGAACGCAGGGAGGGGACCGAGTTGCCTT
GGTAGCTCATGGGCTGGTTGGGGGACAGGACTCCTCGACTGTAGCAGGGTTTCTCCAATCTGTGGGGTAA
CCCGCATCAGAACATGGTGGCAAGTACTTACAAAACATGCGGCTCTCCAGCGGGTTCTTGTCACGCAGAC
ATTCTAGCACCATTGCTTTCAGGAGAAGAGCATGGGCGGGCGCTGACAAGAGTTTAAGAGCTAGAGGGAA
GACGGGGGATGGAAGGAGGGGGTCAGAGAAAGGGAGGGAGCTGCAGCTCACCCTGTTCTCCCACTCCCCA
GTTCATAGGCACAGCCTCCCTTATCGTGTGTGTGCTGGCCATTGTTGACCCCTACAACAACCCCGTCCCC
CGAGGCCTGGAGGCCTTCACCGTGGGCCTGGTGGTCCTGGTCATTGGCACCTCCATGGGCTTCAACTCCG
GCTATGCCGTCAACCCTGCCCGGGACTTTGGCCCCCGCCTTTTTACAGCCCTTGCGGGCTGGGGCTCTGC
AGTCTTCACGTGAGTACAGGGCAGGGCAACAAAGATACACAGGCAAGCCATGCCGTCATCCAGCATGACC
AGAGAGCCACGCAATGGCGCAATGACTAATGAAGAGTGTACTGGAGACCACCACAACGAGAACACGACAC
AACCACAATCAGGTAGAAAAGAAGTGCCACACATCAACGAAAACAAAAACTACCAGGAGAGAACCACAAG
AACAGATGAAGAAGAGTAGTACAGCAGCCGAACAGACCGAGGTAGAACACACGAGAGGAGACAAAAAAAC
CACGCAAA

>DEX0367_002.nt.1
GGSASGWGGCCAAGGCATCTGCCACTGCTGCTGACTGCCACTACCAGCCGGAGCCAGGGCCTGGTCGAGA
CAACTATTTTTGAGGTATCTGAGGCTACCCCACAGCACTCTCTGGCATTGGCAAGACTCGTGGCTGTGAC
CCATGGGCCTCTGAGGAGGCGTTGTAAGTCCGCCCAGCTCCCCACTTGCTGTGTCTCTCACTCAATGCGAA
GGGAACCAAGGTACCAGGGCCAAGTGGGTCCAGACATCCACACGGATTCTGAAAGGGCAGGTCAAGCAGA
CTTTGTCCAGCCATCAGCCACGGCCTCTCGACAGCACCAAGATGGAAGTCAAAGGTCAGCTGATCAGCTC
TCCTACCTTCAATGCCCCAGCTGCCCTGTTCGGGAGAGGCTGCCCCCCAGGTGAAGTCAGAGCGTCTGCGG
GGGCTGCTTGACCGGCAGCGGGACCCTGCAGGAGGCCCTGAGCCTGAAACTTCAGGAGCTCCGCAAAGTGT
GTCTCCAGGAGGCGGAGCTGACTGGCCAGCTGCCCCCTGAGTGCCCACTAGAGCCTGGTGAACGGCCCCA

```
     GTTGGTCCGCCGGCGGCCCCCCACAGCCCGCGCCTACCCTCCACCGCACCCCAACCAAGCACACCACTCC
     TTGTGCCCTGCTGAGGAGCTGGCTCTTGAGGCCCTGGAACGCGAGGTGTCAGTGCAACAGCAGATCGCGG
     CGGCCGCCCGCCGCCTGGCCTTGGCCCCTGATCTGAGCACCGAGCAGCGCCGGCGCCGGCGCCAGGTCCA
     GGCAGATGCACTGAGGAGGCTGCATGAGCTAGAGGAGCAGCTCAGGGATGTCCGGGCCCGCCTTGGCCTC
  5  CCAGTGCTCCCGCTGCCCCAGCCACTGCCACTGTCCACGGGGTCAGTGATCACCACCCAGGGAGTCTGCC
     TGGGCATGCGTCTTGCTCAGCTCAGCCAAGGTGAGCATCCCCTGGTGAGGGTGGGAGAGTGGACACTGGC
     AAATGGACGTGGTAGGGCTGGGATGGGTGACTGGCCGGTGAAAACCGGGAGGTAGGCACAACTCAAGTGG
     CCCAACCCGTGAGAAGAGTTGCGACAAACTGGCTTACCGGAAAGGACAAACTGTGGGGGACTGATGAATA
     TCTGCCAGGACAGTCAGGAACCAAACTGGGGTGGTCTTGGGCTCTGTAAAAGCAGATTCTTGGGACAGG
 10  AACAGGATGAGGCGGAAACTGGCCAGAAGGTGGCCCAAGTAGAGGTGCCCACCCAGTGGTGTAAGGGGGC
     AGCACCCAGGCCTGGGAGTGAGAGGCCCACTTCTCCCCCTCTCACCTGTAGAGGACGTAGTTCTGCACTC
     AGAGAGCAGCTCCCTCTCAGAGTCTGGGGCCAGCCATGACAACGGTGAGGACTCCTATCCCCCAAACCTG
     CCCCCGACTCCTACGTCCCTTTAGCCCATCCCCTTTATTCATCCGACCACATCCTGCATGATCAGCCCAT
     CCTTTGATGCCCAGCCCCTCCTGTGCTGCTTCAGCTTCTCCTTAGCCATGCTCCATCTTCCTATGGGTCC
 15  CCCATCTTCTTAGCCTCACCCCACCCCAGCCCCACAGCCCCAGAACCTACTTGTCAGCCTTAGTTCCTCC
     TCAGCCTGCTTAACCTGTCCCAGCACTTTGCCCCTCCCCATGGCCTTGCCCCCCAGCCTGTCCTGCCCCC
     TCTGGGCCCCCCTCATTGAATTCTATCCCACTTGCTCTCTTCCACTCAGAGGAGCCCCATGGCTGCTTCT
     CTCTGGCCGAGCGCCCCTCACCACCCAAGGCTTGGGACCAGCTGCGGGCAGTATCTGGGGGGAGCCCTGA
     GCGGCGAACCCCATGGAAACCACCTCCATCAGATCTTTATGGGGATCTGAAGAGCCGGCGGAACTCTGTG
 20  GCCAGCCCCACCAGCCCCACACGCTCGCTGCCCAGGAGTGCCTCCAGTTTTGAGGGGCGAAGTGTGCCTG
     CCACCCCTGTCCTCACCCGGGGCGCTGGCCCCCAGCTCTGCAAACCTGAAGGCCTTCATTCTCGTCAGTG
     GTCCGGCAGCCAGGACTCCCAGATGGGCTTCCCCCGGGCGGACCCTGCCTCCGATCGCGCCTCCCTCTTC
     GTAGCTCGCACCCGCCGCAGCAACAGTTCTGAGGCCCTGCTGGTGGACCGGGCCGCTGGTGGGGGAGCTG
     GCTCCCCGCCTGCCCCTCTGGCTCCCTCTGCCTCTGGCCCCCCAGTCTGCAAGAGCAGTGAGGTGCTGTA
 25  TGAGCGCCCCCAACCAACCCCTGCCTTCTCCTCCCGCACAGCAGGCCCCCCAGACCCTCCCCGGGGCCGCC
     CGGCCTAGCTCAGCTGCCCCTGCCTCCCGAGGTGCCCCCCGGCTCCCACCTGTGTGTGGAGACTTCCTCT
     TGGACTATTCCTTGGACCGGGGCCTGCCCCGCAGTGGCGGTGGAACAGGCTGGGGGGAGCTGCCGCCTGC
     AGCTGAGGTCCCAGGACCCCTCTCCCGCCGGGATGGGCTCCTCACCATGCTCCCCGGCCCACCACCTGTG
     TATGCAGCTGACAGCAACAGCCCCCTCCTCCGCACCAAGGACCCCCACACCCGTGCCACCCGCACTAAGC
 30  CCTGTGGCCTGCCCCCAGAGGCTGCCGAAGGCCCTGAGGTGCATCCAAACCCTCTGCTGTGGATGCCCCC
     ACCCACCCGTATCCCCTCGGCTGGTGAACGCAGTGGCCACAAGAACCTGGCTCTGGACGCCTGCGGGGAC
     TGGTACATCCGGAACTCGGGACTGGCTGCGGGGCCCCAGCGCCGGCCTGTGCTCCCTTCCGTGGGCCCGC
     CACACCCACCCTTCCTCCATGCCCGCTGCTATGAGGTGGGGCCAGGCGCTGTACGGGGCCCCCAGCCAGGC
     GCCACTCCCACACTCGAGGAGTTTCACGGCGCCCCCTGTCTCTGGCAGGTATGGGGGGTGCTTTTACTGA
 35  TGGGTAGGGGTCTCGTAAGGCAGATGGCGAAGATATCCAGGCCAGGGAGTGGCTAGTCATGATAGCTAAT
     GAATTGGACCATGAGGAAACTAGCTGCTGTGATGGCACAGGGTCACTCTACTGCACATGACCTGCATTAG
     TCCATGGGGTCCTGGTGGAGGGGATCTTGGGCACTGGTAGCAGCAATTCTTTATCAAGTTATAGGCTGAA
     GATGAGCCTTGAAGCCAGGGTGCCGGGAGGAAGGGACATCTCATGCCCCTTGCTGTTTTCTTCCTTTTTT
     CTCCATGCCCCAGAGCCTGAAAGTGCTGTCCTGTGCCTGCCTCCACCTCTTTAACGAGCCTCTTTTCCTT
 40  TCTTTTTCTGTGTCTTGTCTGTCTTTTCTTCTTCTTGTCTTCCCCGCCCTGTCCTCCGGATTCCTGCTAC
     CCCTTCTAAAGATACTACGCGGACTTCCTGTATCCCCCGGGAGCTGAGCGCTCGTTTAAGTGACCTGACGC
     TAGAGGGGGAGCAGTCCTCCAGTTCTGACACCCAGACCCCGGGGACACTGGTCTGACCCCTTCTGATATG
     TCCCTTGTTGGCCTGGGCACGATTCCAATCTGGGGAGCACACAGCTGACCTCGCTGGGCCCTGGGGTGTG
     GTTGCTCTCAGTCCTGAGCAGAGTGCGCCCAACCTAATCTTCCAAGGCCCCTGGCTCCCCGTAGGCCCAGG
 45  AAGGTGTCTGACACCCTGCTTCTTCTCTCTCAAGCACTGTGCTGGGGACTGGGGGCCCTCAGCTAGCTTAAAAAG
     AGGGGGGATGATGTCATGGGGACCCCAAGCCCCTTCCTCCATTTATGTTTACAGTTGTGACTTAGGTATT
     CACTGTCTTCCTCCAACACTAGGCGTTTTACAAAAGGGAAACTGTGATCTCATCTGGTTGGGTTCATTCC
     TGTTCCCATGCCCAACCAGGTTCCATTCAGGAACCCCCTCCATAAAATGGACCATATCGGGTCTCAGGGC
     CATTTAGGGCAGCCAGGAGACTCCGGTGTGAACAGAAATCCCTGCCACGCATCGCCAGGGCAGTTGGGGC
 50  AGTGGGCTCTCTGCCCACACTTGGAAGGACTGCAGTCTGGGTGGGATGCCTGAAAGAGCCCAACCCCCTC
     TGTGCCCATGGCCTCTGCCCTGACCACCCCCAGTCAGGAGGCCCCACAGGAGGGGCACCCGGTAGATGCC
     AGTGAAATCCTCAGGGGAGGTCTGCCTGAAAGAGCCCAACCCCCTCTGTGCCCATGGCCTCTGCCCTGAC
     CACCCCCAGTCAGGAGGCCCCACAGGAGGGGCACCCGGTAGATGCCAGTGAAATCCTCAGGTGAGGTCTG
     CCTACGGGCCACGGGCCACTCACCACTCACACCTTCCTTGGCTTTCCTTCCACCCTTTTTTTTTTTCTCGA
 55  GACGGAGTCTTGCTCTGTCACCAGGCTGGAGTGCAGCGGCGCAATCTCAGCTCACTGCAACCTCTGCCTC
     CTGGGTTCTCCTGCCTCACCCTCCCGAGTAGCTGGGATTACAGGCACACGCCACCATGCCCGGCTAATTT
     TTGTATTCTTAGTAGAGACAAGGTTTCACCATATTGGCCAGGCTGGTCTTGAACTCTTGACCTCGTGATC
     TGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGCCGTGAGCCACTGCACCCAGCCCCAATCCACCAC
     TTTTTAAGCAAACCCACACAAGTTGTGTTTTCTATGATACCTGTCTGTGATTTTCGGAGCTGGGGGTTCC
 60  CCTACCCCCTTTTCCTGGCGTTAAGCTTTTCTTTTTATACCAGTGGATCTGGACCCAAGACATTACCCAC
```

ACTGGAAGGGGATTTGTATAATAAATGTGTAAACTGAA

>DEX0367_003.nt.1
ACCTCAGAGTGAAACTCACCCAGCTTTAGTAACCAACTCGATTGCATAGACTTTAGATAACCATGTGAAG
GGGATTCTACCATCAGAAAAGAGGCCAAACTTCTATCATCATGGTGGATGTGAAGTGTCTGAGTGACTGT
AAATTGCAGAACCAACTTGAGAAGCTTGGATTTTCACCTGGCCCAATACTACCTTCCACCAGAAAGTTGT
ATGAAAAAAGTTAGTACAGTTGTTGGTCTCACCTCCCTGTGCACCACCTGTGATGAATGGACCCAGAGA
GCTGGATGGAGCGCAGGACAGTGATGACAGCGAAGAGCTTAATATCATTTTGCAAGGAAATATCATACTC
TCAACAGAAAAAAGCAAGAAACTCAAAAAATGGCCTGAGGCTTCCACCACTAAACGCAAAGCTGTAGATA
CCTATTGCTTGGATTATAAGCCTTCCAAGGGAAGAAGGTGGGCTGCAAGAGCACCAAGCACCAGAATCAC
ATATGGGACTATCACCAAAGAGAGAGACTACTGCGCGGAAGACCAGACTATCGAGAGCTGGAGAGAAGAA
GGTTTCCCAGTGGGCTTGAAGCTTGCTGTGCTTGGTATTTTCATCATTGTGGTGTTTGTCTACCTGACTG
TGGAAAATAAGTCGCTGTTTGGTTAAGTAATTTAGGAGCAAAGCAATGCTCCAAGCGAGGCCTCCTGCTT
CAGGAAAGAACCAAAACACTACCCTGAAGGGCCAGCCTAGCCTGCAGCCCTCCCTTGCAGGGAGCCTTCC
CTTGCACTGTGCTGCTCTCACAGATCGGTGTCTGGGCTCAGCCAGGTGGAAGGAACCTGCCTAACCAGGC
ACCTGTGTTAAGAGCATGATGGTTAGGAAATCCCCCAAGTCATGTCAACTCTCATTAAAGGTGCTTCCAT
ATTTGAGCAGGCGTCAAACAAGGAAAAAAAAAGGC

>DEX0367_004.nt.1
AGGCGCCGAGGCCCTCCTGTGCCTGTCACTGGCAATAATCCATGCCACAGAAGACTATTTGCTTTCCCTG
GGCCTAGGTTTGCTGTGAACTTTCAGACTGGCTTCAGTGGAAATGACATTGCCTTCCACTTCAACCCTCG
GTTTGAAGATGGAGGGTACGTGGTGTGCAACACGAGGCAGAACGGAAGCTGGGGGCCCGAGGAGAGGAAG
ACACACATGCCTTTCCAGAAGGGGATGCCCTTTGACCTGCTGCTTCCTGGTGCAGAGCTCAGATTTCAAGG
TGATGGTGAACGGGATCCTCTTCGTGCAGTACTTCCACCGCGTGCCCTTCCACCGTGTGGACACCATCTC
CGTCAATGGCTCTGTGCAGCTGTCCTACATCAGCTTCCAGCCTCCCGGCGTGTGGCCTGCCAACCCGGCT
CCCATTACCCAGACAGTCATCCACACAGTGCAGAGCGCCCCTGGACAGATGTTCTCTACTCCCGCCATCC
CACCTATGATGTACCCCCACCCCGCCTATCCGATGCCTTTCATCACCACCATTCTGGGAGGGCTGTACCC
ATCCAAGTCCATCCTCCTGTCAGGCACTGTCCTGCCCAGTGCTCAGAGGTTCCACATCAACCTGTGCTCT
GGGAACCACATCGCCTTCCACCTGAACCCCCGTTTTGATGAGAATGCTGTGGTCCGCAACACCCAGATCG
ACAACTCCTGGGGGTCTGAGGAGCGAAGTCTGCCCCGAAAAATGCCCTTCGTCCGTGGCCAGAGCTTCTC
AGTGTGGATCTTGTGTGAAGCTCAGTGCCTCAAGGTGGCCGTGGATGGTCAGCACCTGTTTGAATACTAC
CATCGCCTGAGGAACCTGCCCACCATCAACAGACTGGAAGTGGGGGGCGACATCCAGCTGACCCCATGTGC
AGACATAGGCGGCTTCCTGGCCCTGGGGCCGGGGGCTGGGGTGTGGGCAGTCTGGGTCCTCTCATCATC
CCCACTTCCCAGGCCCAGCCTTTCCAACCCTGCCTGGGATCTGGGCTTTAATGCAGAGGCCATGTCCTTG
TCTGGTCCTGCTTCTGGCTACAGCCACCCTGGAACGGAGAAGGCAGCTGACGGGGATTGCCTTCCTCAGC
CGCAGCAGCACCTGGGGCTCCAGCTGCTGGAATCCTACCATCCCAGGAGGCAGGCACAGCCAGGGAGAGG
GGAGGAGTGGGCAGTGAAGATGAAGCCCCATGCTCAGTCCCCTCCCATCCCCCACGCAGCTCCACCCCAG
TCCCAAGCCACCAGCTGTCTGCTCCTGGTGGGAGGTGGCCTCCTCAGCCCCTCCTCTCTGACCTTTAACC
TCACTCTCACCTTGCACCGTGCACCAACCCTTCACCCCTCCTGGAAAGCAGGCCTGATGGCTTCCCACTG
GCCTCCACCACCTGACCAGAGTGTTCTCTTCAGAGGACTGGCTCCTTTCCCAGTGTCCTTAAAATAAAGA
AATGAAAATGCTTGTTGGCACATT

>DEX0367_005.nt.1
CGGGCGCCCAGAGGGCAAATGAGGGGCGGCGGGGTGGCGGGTGGGGGGGAGACGCGGAATTTCCCAGCGC
GGGGGCTCTGGCTTACCCTGCAACCGGGCAGTCTCTTTCTGTTTACGGAGAGAAAGGGGAAATGGAAAAG
TCGGGGAGGCGGTGGCTGGCGTCCGCTGCGCCGCCCCTGGGCAGGCTCAGACGCCGTGAGTCAGGGGCAG
AGCAGGGCGGTCTGAGCGTGCGGGCGACGCGGGTCTCACTCGTCCGCTCCGCTCTGGACTGCGCGCCACG
CTCTGGGGTCCGGCGCCCTGGTTCCTGCTTCTGCCGCTGCCGCCGCCGGATCCCAGTGGCCCGGCGTGCT
CGGCTCCCACAGGCCTGCAGCCAGCATCGCACCGAACCTTCGGGGGGCCGCGGCTGGAGCGCTCGGCCGG
CGTGGGAGCGCCAAGGCCGCAGATGCAATCTTCTTACCGCGAAGAAGCCAGGGGAATAGGTAGCCACATC
TTGTTTGCAGATAAGAAAGGAAGCTAACGCAGTATCTGCAAAGCCAGGAGTCTGACTCAGTACTTTTCTC
ACTCATGCATACAAAGCAGCTAAAAATGACACAGCTTATTTACCATGCCCCTGACACTGCACTGAGCACT
TTATGAGCTTGAACTCTGTTAATCCTCACGACCACCTCATGAGACTCTCCAGAAAGAGCAACAGTAATGG
AGTACATGAGCACTGGAAGTGACAATAAAGAAGAGATTGATTTATTAATTAAACATTTAAATGTGTCTGA
TGTAATAGACATTATGGAAAATCTTTATGCAAGTGAAGAGCCAGCAGTTTATGAACCCAGTCTAATGACC
ATGTGTCAAGACAGTAATCAAAACGATGAGCGTTCTAAGTCTCTGCTGCTTAGTGGCCAAGAGGTACCAT
GGTTGTCATCAGTCAGATATGGAACTGTGGAGGATTTGCTTGCTTTTTGCAAACCATATATCCAACACTGC
AAAGCATTTTTATGGACAACGACCACAGGAATCTGGAATTTTATTAAACATGGTACGTTTCTTTCCGATC
AGTTGGGTGCTATGTGCTCAGCTTTCCTACTGCAGCTTCATTTAATGCAGAGTGGTGTTTGAATTTGGCT
TAGATGGGCTGGCAGGCATGATTTTAAAAAAATGTAATCATCAGTAAGAAGTACTTCCATGTTAAAGATGC

AAAAAAAAAAGTCTTCATTAAAATTTCTATTTAAAGATTCTTGATCATATTTTAGTAGTATTTCTACAAA
TAGGTAATTAAGAAGAATGTATTTGGGCCGGGCGCAGGGGCTCACGCCTGTAATCCGCGCACTTGGAGAG
GCCGAGGTGGGTGGATCACTTGAGGTCAGCAATTAGAGACCAGTCTGGCCAACATGGTGAAACCCCGTCT
CTACTAAAAATACAAAATTAGCCTGGCATGGTGGCGGGCGCCTGTAATTCCAGCTACCTGGGAGGCTGAG
GCACGAGAATCCCTTGAACCTGGGAGGTGGAGGTTGTGATGAGCCGAGATCACACCACTGCACTCCAGCC
TGTGTGACTGAGTGAGACTCTGTCTCAACAACAACAAAGAAGAATGTATTTTAATTTAACAGTTCACTAA
AAAGTTTATTCATCAGTTATGCAGAAAAGAAAGGCTTAATGAGACTCCTTCGTCTGAGAGGGGAATAAAC
ACTTTGTCACT

>DEX0367_006.nt.1

ACCAAACTGAAGTGTAAGGATAAGACCTCTCCTGCTGCTAGTACTAGGAGAGGTCTTATCCTTACACTTC
AGTTTGGTCACCAGATTATGTAAGGATAAGACCTCTCCTAGTACTAGCAGCAGGAATTGTGTGTTCCAGT
AAGTGGTCTCTTGCACGGCACTTTTTTGGGATCAAATGTTAACGTTACTCCCCAGACTCTTCGGGCAAAG
GAATGGCTAGATTCAGAGTAAGAACAACCCTCCCTTTTTTGTAAGTCCCCGTGTTTAGCAGGGAGAAGAA
ATTCTCTAACATGGGTTTGGTTTTGTTGTGTTCTTCATGTGGAAATGTCGCTTAACAAAATATCCAGGCT
TTTGTTTACGTGGAAAAAGCATCCCTTGTAATGATTGCTCATCATACTTAAAAACCTTTTTCAAAGGATT
TTCATGTTCCCAGCTATAAGGACTATTTCCATGACATGTGTTATTGGCAGAATGAGTGTTAAATATGGAG
CATATAGCATGGGGTGACTTTCATTGTCCTAACCTGAGACAGTTTTCCTTATTACTCTGTATTGATCCTG
CTAGTCCAAGAATGGACATGAAGTGAACCTATCGTGGTGACTGGGATAGGAAGGTGCTTGCTATTTTTGC
CAGCACAGCATATTAGTTCCTTTGGAGCCCTCCATTGTCTGAGTCTGCCAGTGATCTGTAGGAAGGCAGCT
GGTCAATAATCATGTAGTACAATGGCTTGGAAAAGAAAAACCATTCAGCTGAGAGTTGCTAACCATGTTCTTT
CAGGCATACTTAGGTATGTCCCTGGGGAAAAAGAAAAACCATTCAGCTGAGAGTTGCTAACCATGTTCTTT
TGGTTAGAAATAATGGTTCATTTTTTGCCCCTGGTTGGAATAGTCTCTAAAAGGCTCTGGTGACTGAATT
GAACATGAGTCCGCATGCTGTTTTCTTTCAAAAGGTATCAAAACGGAAAGCCTCTCTAAGGGGAAGACCT
TTCAACTCCATTCAGAAGACAACATTTAGTAAGGAGGATGGCGGAGGTTACTAGTAATTTTCAGATGTCT
TGGGCTTTTTTCTGCCAACAAAACCCAAAATCAAATTAGAGTTGGTGAAAGCTTTCTTCAGTGTTTTTAGA
AGAGGCCATCTTGAATCTGTAGAATACCATTTACACATCAACTTCACCCTATGCTCATTTCGTATTTTGA
GCTTAAATGGAGTCCTCTGAAGGGGGGGTAGTCGTGATTCTGGTGGCAAAACTAGAACTTTTAACTTGTA
AAATGAAAAATATTAAATGGACCTTTTTGTCAGTTGAGGATTTAGATTGATTCTTTTATCTGAGGAGCGT
ATGTTCCTCAGATGTTGCGTAGAGACCTTTAGGTTTTCATCTACTTTAAGATTGCTCTCTTGGCAATTAG
GGGATTTGGGAAAAGAAGAAAAAAAGATGCCATGTGTTGCTAGTACACAGTTATAGGGATATGGTTTCTAA
TGGTTGAATTTTGAGGAACTCTCCCTAAAGAATGAGTTTTATATCTCCTCAAGGAAATCATGGAAAAATC
TGTTTATTCTTCAGTGAGTCCTTTTGAATTAATGTTCTTAAATTTTTTTTCTAAGTCTGTGTAAGTGCTTA
TGTATAAGTATATAATTGTATAAATATTTATAAATATATTTTATATAATTACGGTTTCATTTCTACCTTGA
GTCAAAGTTCTGTCTTTAGATTGGTGACTGAGTAATACTTACACTTTGGTGTTTTTTCTTAGGTTCTTGA
GGCTTACTTAACTAGCAGCTTCTGATTTATTGAGTGAAAGATGGTTTTCATGTTAATTCCTCAGTTGCAT
CTCTGAACTTGGATAACATACTTGCCGTTTGAAAAATAGAGCTGTATACTGTCAAAGGTGCACTGGAGGG
TAAAACATTTGTTGGTAGTAGACAAGCTCAGAAATCCAAAAATTCMGGGAGGGCACAGGGTAAGAAGAGA
GGTTTGCTCAGCTTTGTGCTCTGATGGCACCCATCTGTACTCCAGCAAAGTCATAGCTGAAGCCCAAATC
GCTCAAACGTGTGAAGTCAATCATTTTTCCTATAGGGCTCTGTTTCCTTTTCTTTACCTTCTAGCTTCTC
TTGATAAAACAGTTGGGAGAACTGCCAACGTGTTCATCAGTGAGAGGGTGTGGTTCTTTCCGATATGTTC
AGCACGTGCATATTCATTTATGAGAAACCAAAGTAACTTCTTCAACTCGGGTTACTGTGGGATGATTAAG
TAGATATAAAGTGATTTAAAAGACAGCTGGTACTTTCACTCAACTTAACATGCAGCTTGGGATGTCTACA
ATACAAGATGTGGTGGGTTTTATTTTTATTTTGAGGCAGGAGTGTAATTATCTTAGGGTTGATATA
AAGGCCCATGGGAAGAGTTACGGATTGTACATGGGAAGAATCCATAAGGAGTATTGCCTTGCTCAAGAAT
TTAATCCGCTGACAGTATGCTTCTTCCAGTTTGCCTTACTTCCTCTGCAAGTTTCTTTGCTCCTTAGGGC
AGAGGTTTTTATTGAATAGATTAAATGAGGTCAGTAAGATTAGTTAAAAATAGTCATTCCTTGTGGGTGTT
GGACACAGCTTGAAAGGAGTGTTTAAAAAAAAAAAAGCTAGGGAGGAAAAGAACAAGAAAGATAGACTTCG
AGAGTGATAAGAGAAAAATACAGGGAGATGGAATGGAAACTACCAGGAAGGCCGGGCGCGGTGGCTCATG
CCTGTAATCCCAACACTTTGGGAGGCGGAGGTGGGCAGATCACCTAAGGTCAGGAGATCGAGGCCAGCCT
GAGCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTACCTGGGTGTGGTGGCGTGCACCTGT
AATCCCAGCTACTCAGAGGCTGAGGCAGGGGAATGGCTTGAACCTGGGAGGCAGAGGTTGCAGTGAGCCA
AGATCGCACCATTACACTCCAGCCTGGGCGACAAAGCGACACTCCATCTCAAATTTAAAAAAAAAAAAAA
AGAAGAAGAAGAAAACTAGTGGGAAAAAAGTGAGAGGAATACTTTTTTGAAATTGGTATCGGAAGGAACT
GGAGAAGAGAAAACAACAGTGCCAAATGAGAAAAGAACAGGAAACTTAATACTAATTGGCATGCACCAGA
TACTTTTGTGTACATTTGCCTCTTCAATCTCCCGAAGAATCGTACAAAATGTATACTTTCTTCCCATGTT
AAGAAAACAGGTATAGAGAACTTATTCAGGACTCATAGCTAGTGAGTGGCAAATCAGAATTGGAATCCAG
AGTCCATGCTGTGGTCTCCTTCAGCAAGAGAATCAAGCTACCCAGATGATTCTCTTCTCATTGTTTGGCT
TTGTAAAGTGTCACTTAGTTTTGTTCCCATAAATATAGAGGAAGTAGTTGGGTTAAAGTTGTGGGATTTG
CACTATGCCTATACATTTTTATTTATGCCATTATTTTGAGAGGCTGATTATTGCTTTTTTAAAATGATGC

ACTGGTGAAGATGTGAAAGTAAAAATTGCCACGTGGCATTATTTGCCAAGAATAAATGAAAAGGGTAAAAA
AACAATTTTTTCTTTTTTTTAAACATTTTTCAGAAAAGATACAATTACCTTTTTTTAATAGGTGAGCGCCAC
CACGCCCGGCTAATTTTTGCATTTTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGGATGGTCTCGATC
TCTTGATCTCGTGTTCCGCCCACCTCGGCCTCCTAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCC

5   GATCTACAATTCTGTTTTCTGTGTCCCATCCATACCAGATTATTGAACCCTCTGTGTGTATGTTTTTAAA
TCTTGTATATAATTCTACTTCCTTGACTTACAAGTCAACAATTAGCTATTGAGTTCTTTATAGAAGTTGG
AAGTATTTTGAAATCTGAGATCTAATTCCTGGCCGGGCGCGGTGGCTCATGCCTGTAATCCCAATATTTT
GGGAGGCCGAGGTGGGTGGATCACCTGAGGTTGGGAGTTCGAGATCAGCCTGGCCAACATGGTGAAACCC
CATCTCTACTGAAAAATACAAAAATTAGCCAGGCATGGTGGCGTGTCTGTAATCCCAGCTAATGGGGAGGC

10   TGAGGCACAAGAATCTCTTGAACCTGGGAGGCGGAGGTTGCAGTGAGCCGAGATCACACCACTGCACTCC
ATCCTGGGCAATGGAGCGAGACTCTGTCTCAAAAAATAATAAATAAGATATGATTCCTGCCTGTTTTATC
CAGTTAGAGAGACACACTACACATGGAGAGCATGTAAATGTTTACCTATTAGTTCAAACTT
AGAAAAAGGGAAGAATATGTTTTAGTTAGCAAAGTCTTAAACAAGACTGAAGTGGCAAGCACTGTAGCTC
ACGTTTTCTTTTGACTCACCTTTTTTCTTTAGCTCCACCTTTTGTAACTTCTAGATTTGGGGCATGAAGTT

15   CTGCCACTAAGGCATAACCAGTACCCAAGTAGGAAAGGATTATTGATCCCTGGAGTAAAGATTCCTAGCT
TGCGGGGATCTCACTATTGATCCCCCGGAGTAAAGATTCCTAGCTCGCAGGAATCTTACTTGCACACTTA
GGGAAAATACCAGAGGGAGCCACTAGCAACCGAAAAATGAAGGAAAGATTTAATTTCTCTACCCAGGTAC
AAATCCAATGCATAGTATAGTATATGTTATATGCAGAAAGGGCACTGGAGGTGTGAGAGCCCTTTGGGTA
ACAGCTCTGCTAGTTGTTACCTTCACACTTCTCTTTCTTCTTAATCTCTGGTTTCTCGTCTGAAAAATGA

20   GATAATTGGACTAGATTAATAGTTCTCAGCCTTTTCTAGAACAACAAATTAGATTATCTCTGAAGTCCTT
TTCCAAAATTTCATGCTTCCAGTCTCTCCCCGTGATTGCGTATTATCCCTAGGCTAATTAAAAGTGTGCT
GGTTAATCAGTTTCTTTACTTTTTTTTTCTCTCTTTTTTTATTCTGTGTTCAGGTATCATTTGACCAACA
...ACTTGCTCTTAATCTTACTGCGATGACTGAAATCTCCAAAGAGGAGAGTATAGTAGTTATAGGTCTAGTA
ATAATAATCTTGTAAACAATTCATTATTTTTTAAAACTTAATGCTGTGTGCCTAATCAGTGTTGAGTATG

25   ACTAAAATGTAAATGAGAACCAAATATGACACTAACGCTGGCACTAAATTTTATTTTTACTTGTGTACTT
AAAGTACTTTACTGGCCTGTTGGGAAATTGATTTGTATAAGATTCACACCTGGTTTTGGCAGAAACAGCT
TTTTAAAGATACAATTAGGCTGACTTAAAAATGTTTTTTATTCCACAGAATGTATCTTTATTATGTCATT
CTTACCGTCTCTCTGCCTTTTTTGACCATCATGTAACTACGCAGTAGTAAATCATGGGGTGCAGGCTGAT
GAATTAGTAGTCTTTGTAAAACTAAAATGTAATGAGCATAGGAAGAGACAGCTCTTCTGGTGAACATGGC

30   AGTTTTCCGGTAATACCAGCCTTGATGAATAGGCTAAAGACAGCTGTGTACTATGGACATTCAACTTCAC
TGGATTGCCTGGCTTAGCTTTTCAGTTATTGTCTTGCCGTGAAACATGGTGTTGCACTGTACACATTTTA
GAATAGTGGCAAGAATGTTCTTTAGAACTGCAGATAACCCAACAGAGAATACCATACCAATATATCGGTT
TCTCCTGTTACATAGAGATTTGGTATTTCAGTGATGGTTATGGTTTTTGACTCACCTGTGTGTCAGTCCT
GCTAGAGAGACAGTATAGAATCAAGAATTTCTTGATGTCTTTTTAAAAAATAGAGATTATAATGGCCTTGC

35   TAGTTGTCCTGTGAAGTGACAGACCCTTAATTAGAAATCTTTTTAATCCTGCTTTCCATTTTACTTCTCCA
CCATTTATTTTTAACATTCATTCTCTGAAAATAGTTATGATTTTTAGTGTAATTGGCTATGTTTAAGAAA
TTCCATTTCAAAGTCTTAATATGTAACAATCTTCATTTTTTATAAAATTGAAGGACATTTTTGAGGCAAA
CTTACTTTTATAGCTCATTTTCTCCCTAGTTAAGGAGAATTCTCTGGTTTCATATAGAGAATGGCTAGCT
CTGTGATACCCCACCCTCTTTGTGGCCTACCAGTTGCTGTTGTGTTGCTTAAAATTGCAGATATTCTCAT

40   AAACATGATGATTTTTGGCACAGGCCTTTCTTTGCTATTGATTTCAAAATAAAGTTGGGCAATCCAATTT
GAACTACTTTAATAAACATAATTTAGCATTCCTGTAATGAGAAAGTATTTTCAAAAGATAACCAGATTTA
TTTCTAGGATTAGAGATGGCAAATACCAATGATAGTATTTTTCCCAAGAGCAATCACAGATTAATCTATA
AATATGAGAAGTCGCATTATTGTACTTAATTCTTTGGCAACAAACAGCAAGTTTTTATTGTTTAATTGTA
TTTCCTCTCTAGAGATCACAATACTCTGGGTATTTTATATCCCTTCTAAAGGTCATTTGGGTTTTCAAAT

45   GGGAAGAATAGTCAAGCTAAGCTGGACTAAACCTAAGTAATATTTTTTCTCATCAAAAGAAGTTATTAAT
CTAACTGGGTTAGCATGAGTCATTCATTTTTTAGACATGATAAATGGGAAACATATCAAATCATTAGATA
AATTCTGACCTGGAATTGAATTCCCCCTTTCTTAAAATCTTTTTAATTTGCTTTTTTTCATAGACTCAGTA
TGGAACTCTTAACTGATAAGGGAGAGATTCTTGATCTGGAACCGTTCCCTGCCATTCTCCTTTTTTCTCT
CTGTCTAGGTAGCTGGTTCCATAGTGCCCGTCACGAGGGTCCATTTCAGTTTGATGACATTAGACTGCTT

50   ACACTCAGCTGGATGCCTTGCTGTTTGCAGCAACATGATTTTACAGTATGCTTTTCTCAAAGCTTTGCAT
TCTTAATGGAGATATCAAATGGTGTAATTCCAAATATAAATATGTTTATGACACTAATCATATGCTTTTA
ACAATAACTTTTTGATAACTTATTGCCCTGTAAGTTAAACCTTACACGGAGTGACTGTTGCCATCAGAAA
ATCCCCTTTCACCTGCTAGAGAGAGTGAAAGTAGTGCGTGGCCTCCACTTTTCAGTTAACGAGTAGCAAG
CTTCTAGTAGGGGCTGCTTATCTGACCCATGTGGGGTCTTGGGCCTTGTTACTTTCCTGGGTTCGTCCCT

55   TCAGCTGGAAAGGCTGTTGAAAACACTTGCCAGGAAACAGTAAAGCTGTGAGAGAATCTTCTAGTCTTAG
ATATAGAGAAATGAGTACCAGTTGATGCTAACACGATACTTAGACTTTGAGGGGCTTCCATCCAAGCTTT
GGCATGAATCTGTTAGACACGGTTTTCTTCACTCCTCTGATAATCTTTTTTTTCCAGAATGGTTTTTTGTT
GTTGTTGTTGTTAACAAATTCTAAACATCCAGTGTTACTTTTTGTTTTGTTTGTTTTGTTTTGTT
TTGTTTTGTTTGCCCTTCAGGAAAGTCCTTTTTCGAGTAGGTATTGATTGTTCCCTGGAAACTCTCAGCA

60   GTGTCTGTGCAGGCTCTGTGCATGCTTTGTATGAGTTCGGCTTGAACAATGCATTTGAAGTCACCTGGGA

```
     TGTCCAGTTCTGGCATGTCTTCATTGATTGTGTTTTTAAACATGTTTCATGTTTCATGTCATTTTCCAAA
     CCTCACTTTACATCTTACTCAGAGAAGTTAATCAAGGAATGAAATTTCTAGGGGGATAGGGATGATGAGG
     GTGGGTTGGGGGCTTGAGTGAAGTCAACTTGGGGTATTTGCTTTAAAGTGTTTTCTAAAGCAGTTCCTAA
     CAGTATTTAGAGATTCCCTGGGATGTTTGTGGCTCAACTTATTCTGGACAAAGTGTGTGTGGGGGTGGTC
5    TTACTGAGATTTGCATTCTTAATGCAGGACAGGCTCTAAATTTCATCTGTACTCTAAATTTGTGATTGAA
     TCCAAGAAGATAACAGAGACAGTGCTCCTGTTGTAATGTTTCTGGCAAGTGCTCCCTAAAATGCACATCG
     AATTCTGTTTTCTGGGCCTTTTCTCCAATGGTGCTAGGAGATACCGTTGATTTCTGCAGCTCTTCTCAGT
     GGTGGGAAGAAGTCTTTGGGATTGTTGAGCAAGGGGCAGCTGGACCATCCACTAAATTTTTTTTGTTCAAG
     ACACATTAGAGACCCTCCTGTATATCTAGTAAGTCATAATAAAGGTGCTTGGGAAAGCCTTAAATTTGAA
10   GACACATGGAGGCGGTAGAAAATTAAACTTGTAAGAGGAGAAAAACATGCCATTAGGTAACGCAGAGTTG
     TAACTACTGGCTAAGACTCAATGGAACTTCCACTTGCTCTAAAACCAGGGAAGGGAGTGGGACGATAAGT
     CTTTTGAAGACATCAGCTCACTGTGCTGAGAGAGGGACCAAACTCAAGGAAACCTCTGATCTATACATTC
     AACTGCTGCATTTTTTTATAAATACATGTAAATGTCCTGTTGTAATATTTGATCTTGGAAATAAAAACAA
     AAACTTTTCAGTAGAAAAAAAACAAAAAAAAAAAAATTTKGGGGGCCGGCGCCAAGAAGTTTCAACCACA
15   TGTGTGGGGCCGGCCCAGGTGAAGGAAATGGCAAGTGGCCAAGAACCGGCCAGAAGGGGGGTTTCCCAAA
     AAACGCGCGGGCGCGCGTTTGA
```

>DEX0367_006.nt.2
```
20   TTCTGGTGGCAAAACTAGAATCACGACTACCCCCCCCTTCAGAGGACTCCATTTAAGGGGGGGTAGTCGTG
     ATTCTGGTGGCAAAACTAGAACTTTTAACTTGTAAAATGAAAAATATTAAATGGACCTTTTTGTCAGTTG
     AGGATTTAGATTGATTCTTTTATCTGAGGAGCGTATGTTCCTCAGATGTTGCGTAGAGACCTTTAGGTTT
     TCATCTACTTTAAGATTGCTCTCTTGGCAATTAGGGGATTTGGGAAAAGAAGAAAAAAAGATGCCATGTG
     TTGCTAGTACACAGTTATAGGATATGGTTTCTAATGGTTGAATTTTGAGGAACTCTCCCTAAAGAATGAG
     TTTTATATCTCCTCAAGGAAATCATGGAAAAATCTGTTTATTCTTCAGTGAGTCCTTTTGAATTAATGTT
25   CTTAAATTTTTTTTCTAAGTCTGTGTAAGTGCTTATGTATAAGTATATAATTGTATAAATATTTATAAATA
     TATTTATATAATTACGGTTTCATTTCTACCTTGAGTCAAAGTTCTGTCTTTAGATTGGTGACTGAGTAAT
     ACTTACACTTTGGTGTTTTTTCTTAGGTTCTTGAGGCTTACTTAACTAGCAGCTTCTGATTTATTGAGTG
     AAAGATGGTTTTCATGTTAATTCCTCAGTTGCATCTCTGAACTTGGATAACATACTTGCCGTTTGAAAAA
     TAGAGCTGTATACTGTCAAAGGTGCACTGGAGGGTAAAACATTTGTTGGTAGTAGACAAGCTCAGAAATC
30   CAAAAATTCMGGGAGGGCACAGGGTAAGAAGAGAGGTTTGCTCAGCTTTGTGCTCTGATGGCACCCATCT
     GTACTCCAGCAAAGTCATAGCTGAAGCCCAAATCGCTCAAACGTGTGAAGTCAATCATTTTTCCTATAGG
     GCTCTGTTTCCTTTTCTTTACCTTCTAGCTTCTCTTGATAAAACAGTTGGGAGAACTGCCAACGTGTTCA
     TCAGTGAGAGGGTGTGGTTCTTTCCGATATGTTCAGCACGTGCATATTCATTTATGAGAAACCAAAGTAA
     CTTCTTCAACTCGGGTTACTGTGGGATGATTAAGTAGATATAAAGTGATTTAAAAGACAGCATGGTACTT
35   TCACTCAGCTTAACATGCAGCTGGGATGTCTACAATACAAGATGTGGTGGGTTTTATTTTATTTTATTTT
     GAGGCAGGAGTGTAATTATCTTAGGGTTGATATAAAGGCCCATGGGAAGAGTTACGGATTGTACATGGGA
     AGAATCCATAAGGAGTATTGCCTTGCTCAAGAATTTAATCCGCTGACAGTATGCTTCTTCCAGTTTGCCT
     TACTTCCTCTGCAAGTTTCTTTGCTCCTTAGGGCAGAGGTTTTATTGAATAGATTAAATGAGGTCAGTAA
     GATTAGTTAAAAAATAGTCATTCCTTGTGGGTGTTGGACACAGCTTGAAAGGAGTGTTTAAAAAAAAAAAG
40   CTAGGGAGGAAAAGAACAAGAAAGATAGACTTCGAGAGTGATAAGAGAAAAATACAGGGAGATGGAATGG
     AAACTACCAGGAAGGCCGGGCGCGGTGGCTCATGCCTGTAATCCCAACACTTTGGGAGGCCGGAGGTGGGC
     AGATCACCTAAGGTCAGGAGATCGAGGCCAGCCTGAGCAACATGGTGAAACCCCGTCTCTACTAAAAATA
     CAAAAATTACCTGGGTGTGGTGGCGTGCACCTGTAATCCCAGCTACTCAGAGGCTGAGGCAGGGGAATGG
45   CTTGAACCTGGGAGGCAGAGGTTGCAGTGAGCCAAGATCGCACCATTACACTCCAGCCTGGGCGACAAAG
     CGACACTCCATCTCAAATTTAAAAAAAAAAAAAAAGAAGAAGAAGAAAACTAGTGGGAAAAAAGTGAGAG
     GAATACTTTTTTGAAATTGGTATCGGAAGGAACTGGAGAAGAGAAAACAACAGTGCCAAATGAGAAAAGA
     ACAGGAAACTTAATACTAATTGGCATGCACCAGATACTTTTGTGTACATTTGCCTCTTCAATCTCCCGAA
     GAATCGTACAAAATGTATACTTTCTTCCCATGTTAAGAAAACAGGTATAGAGAACTTATTCAGGACTCAT
     AGCTAGTGAGTGGCAAATCAGAATTGGAATCCAGAGTCCATGCTGTGGTCTCCTTCAGCAAGAGAATCAA
50   GCTACCCAGATGATTCTCTTCTCATTGTTTGGCTTTGTAAAGTGTCACTTAGTTTTGTTCCCATAAATAT
     AGAGGAAGTAGTTGGGTTAAAGTTGTGGGATTTGCACTATGCCTATACATTTTTATTTATGCCATTATTT
     TGAGAGGCTGATTATTGCTTTTTTAAAATGATGCACTGGTGAAGATGTGAAAGTAAAATTGCCACGTGGC
     ATTATTTGCCAAGAATAAATGAAAAGGGTAAAAAAACAATTTTTTCTTTTTTTAAACATTTTTCAGAAAA
     GATACAATTACCTTTTTTAATAGGTGAGCGCCACCACGCCCGGCTAATTTTTGCATTTTTAGTAGAGACA
55   GGGTTTCACCATGTTGGCCAGGATGGTCTCGATCTCTTGATCTCGTGTTCCGCCCACCTCGGCCTCCTAA
     AGTGCTGGGATTACAGGCATGAGCCACTGCGCCCGATCTACAATTCTGTTTTCTGTGTCCCATCCATACC
     AGATTATTGAACCCTCTGTGTGTATGTTTTTAAATCTTGTATATAATTCTACTTCCTTGACTTACAAGTC
     AACAATTAGCTATTGAGTTCTTTATAGAAGTTGGAAGTATTTTGAAATCTGAGATCTAATTCCTGGCCGG
     GCGCGGTGGCTCATGCCTGTAATCCCAATATTTTGGGAGGCCGAGGTGGGTGGATCACCTGAGGTTGGGA
60   GTTCGAGATCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTGAAAAATACAAAAATTAGCCAGGCAT
```

```
      GGTGGCTGTCTGTAATCCCAGCTAATGGGGAGGCTGAGGCACAAGAATCTCTTGAACCTGGGAGGCGGAG
      GTTGCAGTGAGCCGAGATCACACCACTGCACTCCATCCTGGGCAATGGAGCGAGACTCTGTCTCAAAAAA
      TAATAAATAAGATATGATTCCTGCCTGTTTTATCCAGTTAGAGAGACACACTACACATGAAATAATTGGA
      GAGCATGTAAATGTTTACCTATTAGTTCAAACTTAGAAAAAGGGAAGAATATGTTTTAGTTAGCAAAGTC
5     TTAAACAAGACTGAAGTGGCAAGCACTGTAGCTCACGTTTTCTTTTGACTCACCTTTTTCTTTAGCTCCA
      CCTTTTGTAACTTCTAGATTTGGGGCATGAAGTTCTGCCACTAAGGCATAACCAGTACCCAAGTAGGAAA
      GGATTATTGATCCCTGGAGTAAAGATTCCTAGCTTGCGGGGATCTCACTATTGATCCCCCGGAGTAAAGA
      TTCCTAGCTCGCAGGAATCTTACTTGCACACTTAGGGAAAATACCAGAGGGAGCCACTAGCAACCGAAAA
      ATGAAGGAAAGATTTAATTTCTCTACCCAGGTACAAATCCAATGCATAGTATAGTATATGTTATATGCAG
10    AAAGGGCACTGGAGGTGTGAGAGCCCTTTGGGTAACAGCTCTGCTAGTTGTTACCTTCACACTTCTCTTT
      CTTCTTAATCTCTGGTTTCTCGTCTGAAAAATGAGATAATTGGACTAGATTAATAGTTCTCAGCCTTTTC
      TAGAACAACAAATTAGATTATCTCTGAAGTCCTTTTCCAAAATTTCATGCTTCCAGTCTCTCCCCGTGAT
      TGCGTATTATCCCTAGGCTAATTAAAAGTGTGCTGGTTAATCAGTTTCTTTACTTTTTTTTCTCTCTTTT
      TTTTATTCTGTGTTCAGGTATCATTTGACCAACAACTTGCTCTTAATCTTACTGCGATGACTGAAATCTC
15    CAAAGAGGAGAGTATAGTAGTTATAGGTCTAGTAATAATAATCTTGTAAACAATTCATTATTTTTTAAAA
      CTTAATGCTGTGTGCCTAATCAGTGTTGAGTATGACTAAAATGTAAATGAGAACCAAATATGACACTAAC
      GCTGGCACTAAATTTTATTTTTACTTGTGTACTTAAAGTACTTTACTGGCCTGTTGGGAAATTGATTTGT
      ATAAGATTCACACCTGGTTTTGGCAGAAACAGCTTTTTAAAGATACAATTAGGCTGACTTAAAAATGTTT
      TTTATTCCACAGAATGTATCTTTATTATGTCATTCTTACCGTCTCTCTGCCTTTTTTGACCATCATGTAA
20    CTACGCAGTAGTAAATCATGGGGTGCAGGCTGATGAATTAGTAGTCTTTGTAAAACTAAAATGTAATGAG
      CATAGGAAGAGACAGCTCTTCTGGTGAACATGGCAGTTTTCCGGTAATACCCAGCCTTGATGAATAGGCTA
      AAGACAGCTGTGTACTATGGACATTCAACTTCACTGGATTGCCTGGCTTAGCTTTTCAGTTATTGTCTTG
      CCGTGAAACATGGTGTTGCACTGTACACATTTTAGAATAGTGGCAAGAATGTTCTTTAGAACTGCAGATA
      ACCCAACAGAGAATACCATACCAATATATCGGTTTCTCCTGTTACATAGAGATTTGGTATTTCAGTGATG
25    GTTATGGTTTTTGACTCACCTGTGTGTCAGTCCTGCTAGAGAGACAGTATAGAATCAAGAATTTCTTGAT
      GTCTTTTTAAAAATAGAGATTATAATGGCCTTGCTAGTTGTCCTGTGAAGTGACAGACCCTTAATTAGAA
      ATCTTTTAATCCTGCTTTCCATTTTACTTCTCCACCATTTATTTTTAACATTCATTCTCTGAAAATAGTT
      ATGATTTTTAGTGTAATTGGCTATGTTTAAGAAATTCCATTTCAAAGTCTTAATATGTAACAATCTTCAT
      TTTTTATAAAATTGAAGGACATTTTTGAGGCAAACTTACTTTTATAGCTCATTTTCTCCCTAGTTAAGGA
30    GAATTCTCTGGTTTCATATAGAGAATGGCTAGCTCTGTGATACCCCACCCTCTTTGTGGCCTACCAGTTG
      CTGTTGTGTTGCTTAAAATTGCAGATATTCTCATAAACATGATGATTTTTGGCACAGGCCTTTCTTTGCT
      ATTGATTTCAAAATAAAGTTGGGCAATCCAATTTGAACTACTTTAATAAACATAATTTAGCATTCCTGTA
      ATGAGAAAGTATTTTCAAAAGATAACCAGATTTATTTCTAGGATTAGAGATGGCAAATACCAATGATAGT
      ATTTTTCCCAAGAGCAATCACAGATTAATCTATAAATATGAGAAGTCGCATTATTGTACTTAATTCTTTG
35    GCAACAAACAGCAAGTTTTTATTGTTTAATTGTATTTCCTCTCTAGAGATCACAATACTCTGGGTATTTT
      ATATCCCTTCTAAAGGTCATTTGGGTTTTTCAAATGGGAAGAATAGTCAAGCTAAGCTGGACTAAACCTAA
      GTAATATTTTTTCTCATCAAAAGAAGTTATTAATCTAACTGGGTTAGCATGAGTCATTCATTTTTTAGAC
      ATGATAAATGGGAAACATATCAAATCATTAGATAAATTCTGACCTGGAATTGAATTCCCCCTTTCTTAAA
      ATCTTTTTAATTTGCTTTTTTCATAGACTCAGTATGGAACTCTTAACTGATAAGGGAGAGATTCTTGATC
40    TGGAACCGTTCCCTGCCATTCTCCTTTTTTCTCTCTGTCTAGGTAGCTGGTTCCATAGTGCCCGTCACGA
      GGGTCCATTTCAGTTTGATGACATTAGACTGCTTACACTCAGCTGGATGCCTTGCTGTTTGCAGCAACAT
      GATTTTACAGTATGCTTTTCTCAAAGCTTTGCATTCTTAATGGAGATATCAAATGGTGTAATTCCAAATA
      TAAATATGTTTATGACACTAATCATATGCTTTTAACAATAACTTTTTGATAACTTATTGCCCTGTAAGTT
      AAACCTTACACGGAGTGACTGTTGCCATCAGAAAATCCCCTTTCACCTGCTAGAGAGAGTGAAAGTAGTG
45    CGTGGCCTCCACTTTTCAGTTAACGAGTAGCAAGCTTCTAGTAGGGGCTGCTTATCTGACCCATGTGGGG
      TCTTGGGCCTTGTTACTTTCCTGGGTTCGTCCCTTCAGCTGGAAAGGCTGTTGAAAACACTTGCCAGGAA
      ACAGTAAAGCTGTGAGAGAATCTTCTAGTCTTAGATATAGAGAAATGAGTACCAGTTGATGCTAACACGA
      TACTTAGACTTTGAGGGGCTTCCATCCAAGCTTTGGCATGAATCTGTTAGACACGGTTTTCTTCACTCCT
      CTGATAATCTTTTTTTCCAGAATGGTTTTTTGTTGTTGTTGTTGTTAACAAATTCTAAACATCCAGTGTT
50    ACTTTTTGTTTGTTTGTTTTGTTTTGTTTTGTTTTGTTTGCCCTTCAGGAAAGTCCTTTTTCGA
      GTAGGTATTGATTGTTCCCTGGAAACTCTCAGCAGTGTCTGTGCAGGCTCTGTGCATGCTTTGTATGAGT
      TCGGCTTGAACAATGCATTTGAAGTCACCTGGGATGTCCAGTTCTGGCATGTCTTCATTGATTGTGTTTT
      TAAACATGTTTCATGTTTCATGTCATTTTCCAAACCTCACTTTACATCTTACTCAGAGAAGTTAATCAAG
      GAATGAAATTTCTAGGGGGATAGGGATGATGAGGGTGGGTTGGGGGCTTGAGTGAAGTCAACTTGGGGTA
55    TTTGCTTTAAAGTGTTTTCTAAAGCAGTTCCTAACAGTATTTAGAGATTCCCTGGGATGTTTGTGGCTCA
      ACTTATTCTGGACAAAGTGTGTGTGGGGGTGGTCTTACTGAGATTTGCATTCTTAATGCAGGACAGGCTC
      TAAATTTCATCTGTACTCTAAATTTGTGATTGAATCCAAGAAGATAACAGAGACAGTGCTCCTGTTGTAA
      TGTTTCTGGCAAGTGCTCCCTAAAATGCACATCGAATTCTGTTTTCTGGGCCTTTTCTCCAATGGTGCTA
      GGAGATACCGTTGATTTCTGCAGCTCTTCTCAGTGGTGGGAAGAAGTCTTTGGGATTGTTGAGCAAGGGG
60    CAGCTGGACCATCCACTAAATTTTTTTGTTCAAGACACATTAGAGACCCTCCTGTATATCTAGTAAGTCA
```

```
TAATAAAGGTGCTTGGGAAAGCCTTAAATTTGAAGACACATGGAGGCGGTAGAAAATTAAACTTGTAAGA
GGAGAAAAACATGCCATTAGGTAACGCAGAGTTGTAACTACTGGCTAAGACTCAATGGAACTTCCACTTG
CTCTAAAACCAGGGAAGGGAGTGGGACGATAAGTCTTTTGAAGACATCAGCTCACTGTGCTGAGAGAGGG
ACCAAACTCAAGGAAACCTCTGATCTATACATTCAACTGCTGCATTTTTTTATAAATACATGTAAATGTC
```
5
```
CTGTTGTAATATTTGATCTTGGAAATAAAAACAAAAACTTTTTCAGTAGAAAAAAAAACAAAAAAAAAAAA
TTTKGGGGGCCGGCGCCAAGAAGTTTCAACCACATGTGTGGGGCCGGCCCAGGTGAAGGAAATGGCAAGT
GGCCAAGAACCGGCCAGAAGGGGGGTTTCCCAAAAAACGCGCGGGCGCGCGTTTGA
```

>DEX0367_007.nt.1
10
```
GACCTCTCCTCATCTGTGCATGGATGAGTGGCCGACTTTGGAGCCCAGGCTGTTACTTCCTGGTCTGGTG
GTGAATCCTCCATAGTCTGAGAGTAAGATCCTTGATACTGGCTCAGCATGGAACATCTGGCACACAGTAT
GCACTGAGGAAATACTTGTTGGAATAATCAGTGAATCATAGATGAAAACTTAACCTTGGAATTAATTATG
AGACTGCTCAGAGGAAGAGAATGGGAGACAAAGGACCTGGTGATTAGACCCCCAAGACACTGGGCTGTCT
GCTTGTGTCTCGGGTGGAACAGGCCCAGCGAGAGTCTTTAGGGCCAGAACTCAAGAATTTATTGAGCCCT
```
15
```
TGTTCTAGGCACTTGGGATTCACCAGTATACAATGGAGACAAAAATCCCTGCCCTGGAGCAGCTTACATT
CTAGCATGGCAAACAGGCAGTAAACAGCCCATTCTGGCTGCTGTATTGAGAAGAGAATGTGGTGGACAGA
TATAGAAGCATGGAAACCTGATAGGGCTATTGCAATCACTCAGAAAAGAGGCGATGGCAGCTTGGACCTG
TTGGAAGCAGTAGAGTGCTTCCAGGGAGGAGAAAGGACCTGAAGGTTAATTTGATCACCAATGGCCAATG
ATGTAATCAGTTATGCCTATGTAATGAAGCTTCCGTAAAAACCAAAAGGATTGGGTTTGATTGAGCTTCT
```
20
```
GGATAGCTAAACATGTGGAAGTTCCTGGAGGATGGTGGTCCTAAACAGGGTTACCAGAAAATTTAAAGTT
ACGCCAGTCACTCCCATTCTATTGCTGAGAGAAAATACAGAGCGGTGAAGAGCTTACAGTCGGGTGAGCA
AGATACGCCCCCAATGTTAGCTAATGATGTAAGGCATCAACAAGAAATGTGGGGGTTCAGGAAAGTGGAA
GGTGGGGTCGTCCAAAGCCTGGGAAAGTCTTCAGTTGAAGGCGAGACAGATGGGACTATTAGTGAGTTTC
GAGAGATCCAACGCCTCGCCGCCTTCGCTTCATTTTTGTCTCACGCCCCACCCTTGAACGCGCGGCGCCT
```
25
```
TTTGACCCCGCCGCCGCGCCGTAGACCGCGCTGCACGCCGGCGGCAGCCATGGCGGATGTGTCAGAGAGG
ACACTGCAGTTGTCCGTGCTAGTAGCCTTCGCTTCTGGAGTACTCCTGGGCTGGCAGGCGAACCGACTGC
GGAGGCGCTACTTGGACTGGAGGAAAAGGAGGCTGCAGGACAAGCTGGCGGCGACGCAGAAGAAGCTGGA
CCTGGCCTGAGACTCTGCGCCTTCCGCCCCATTCTGTCCCCCTCATGGCCACCTTGCCATGCTCCGCGCCG
GAGCCCCGGTCCGCCGGCGCCCAGAACCAGGCTTGCCACACAGTCCCCGCCTGCCATGGCCGGCTCGGCCG
```
30
```
TGGAGTGTTGCTTGTTGAAGATGCATATAGACTACCGGAAAGAAATGACTTCTTCTTGCACTTTATGAAT
CAATTTTATTTTTAAAATAAAAAAATTAAACATCAAAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAAGGAA
AAAGAAAAAAAAAAAAAAAAAGAAACAAGACAAAAACAGAGAGCGGCAAAAAAAGAACACAGAGAGGAG
ACACACAAAGACGAACAAACCACGAGAAAGAGCGACCAAGAAGCGCACAAAAAAAAGAAAAGAGGAGGCA
CAGAAACAACCGCGAGGAAAAAGAACAACAGCAGACCGAGAGCGCCCGGAGCGGGAAAACACACACGAAG
```
35
```
TGAGGACAGAGAAGACGACAGCCGGACACAAACAAAAAGAAAGGAAAAGAGACACGAGCGGAGGGCGACA
CAAAAAACAAAAGACAGAACAAACGACGCGAAGACACCCGAGACAAAAGCGACCAGGAGAGAGAAGCGAA
CGAGCGCCCACCCCGACAACACAAGACAAGA
```

>DEX0367_007.nt.2
40
```
GACCTCTCCTCATCTGTGCATGGATGAGTGGCCGACTTTGGAGCCCAGGCTGTTACTTCCTGGTCTGGTG
GTGAATCCTCCATAGTCTGAGAGTAAGATCCTTGATACTGGCTCAGCATGGAACATCTGGCACACAGTAT
GCACTGAGGAAATACTTGTTGGAATAATCAGTGAATCATAGATGAAAACTTAACCTTGGAATTAATTATG
AGACTGCTCAGAGGAAGAGAATGGGAGACAAAGGACCTGGTGATTAGACCCCCAAGACACTGGGCTGTCT
GCTTGTGTCTCGGGTGGAACAGGCCCAGCGAGAGTCTTTAGGGCCAGAACTCAAGAATTTATTGAGCCCT
```
45
```
TGTTCTAGGCACTTGGGATTCACCAGTATACAATGGAGACAAAAATCCCTGCCCTGGAGCAGCTTACATT
CTAGCATGGCAAACAGGCAGTAAACAGCCCATTCTGGCTGCTGTATTGAGAAGAGAATGTGGTGGACAGA
TATAGAAGCATGGAAACCTGATAGGGCTATTGCAATCACTCAGAAAAGAGGCGATGGCAGCTTGGACCTG
TTGGAAGCAGTAGAGTGCTTCCAGGGAGGAGAAAGGACCTGAAGGTTAATTTGATCACCAATGGCCAATG
ATGTAATCAGTTATGCCTATGTAATGAAGCTTCCGTAAAAACCAAAAGGATTGGGTTTGATTGAGCTTCT
```
50
```
GGATAGCTAAACATGTGGAAGTTCCTGGAGGATGGTGGTCCTAAACAGGGTTACCAGAAAATTTAAAGTT
ACGCCAGTCACTCCCATTCTATTGCTGAGAGAAAATACAGAGCGGTGAAGAGCTTACAGTCGGGTGAGCA
AGATACGCCCCCAATGTTAGCTAATGATGTAAGGCATCAACAAGAAATGTGGGGGTTCAGGAAAGTGGAA
GGTGGGGTCGTCCAAAGCCTGGGAAAGTCTTCAGTTGAAGGCGAGACAGATGGGACTATTAGTGAGTTTC
GAGAGATCCAACGCCTCGCCGCCGTAGACCGCGCTGCACGCCGGCGGCAGCCATGGCGGATGTGTCAGAGAGG
```
55
```
TTTGACCCCGCCGCCGCGCCGTAGACCGCGCTGCACGCCGGCGGCAGCCATGGCGGATGTGTCAGAGAGG
ACACTGCAGTTGTCCGTGCTAGTAGCCTTCGCTTCTGGAGTACTCCTGGGCTGGCAGGCGAACCGACTGC
GGAGGCGCTACTTGGACTGGAGGAAAAGGAGGCTGCAGGACAAGCTGGCGGCGACGCAGAAGAAGCTGGA
CCTGGCCTGAGACTCTGCGCCTTCCGCCCCATTCTGAGCCCCGGTCCGCCGGCGCCCAGAACCAGGCTTG
CCACACAGTCCCCGCCTGCCATGGCCGGCTCTTCCTGGAGTGTTGCTTGTTGAAGATGCATATAGACTAC
```
60
```
CGGAAAGAAATGACTTCTTCTTGCACTTTATGAATCAATTTTATTTTTAAAATAAAAAAATTAAACATC
```

176

>DEX0367_008.nt.1
TTCCTGGTGGTGATGGTGTGCACTTTTATTCAACTGGTCTCAAGTCAGTGTACAGGTAAGCCCTGGCTGC
CTCCACCCACTCCCAGGGAGACCAAAAGCCTTCATACATCTCAAGTTGGAGGACAAAGGAAGGGAGAGGG
GTGGAGGGACAGAAAGGAAGAGGAAGGATGGAGGAAGGAAAAAGGAGGAGAAGCAAACGGGGGGGGGCTGC
CCTTGCCTCTGAGAACCTGCTACACTGTCTCCTTTGGTAGCAGCCTGCCCCCTGTGCCTCTAACCTACACG
TGACTTGGGTCAAATAAAACTTCCAGGAGGATCCAAACCTGAAGGACAGTGGGAACCCCAGAATGCACTT
TGAGGCTTCTCTAGCAGTGGACGAGCGCCTGTCTTCTTAGTCCCTGACCCTGGATCTTTCTGAAGGTCTC
CTCATCACCCAGCTGCCAGGGAACCTGGCTGCCCCTGAAGCAGCTCACACTGGTCTGTCCCCTCTGCAGC
ACCACCCATGCTGCCCCTCAGCTCTTCCCGCTGCTCTCACTCCTGAACACTTAAGGAGATGGCAGGGGAG
GAGCTGCCGAACCAGAGGAGCACCCCACACACTGCCAACAGGGCTTGACCCAGGAGCGGGGGCTTGAAAA
GCTCCAGCTCCGTTTTTAGAGGAAGAAGAATGCCGGATTACAACACAGCACCAGCAATCTGGATATTCGT
TCCCTGGCTGTTCTTTAACCAGCCAGGAAATGCCCTGTGCAAGGGCTGGAGGCCTGGGGCTGCAGACTCC
CAATCTCAATGGGCACCCCAGGGCTGAGCCACCTGAGGGCACTGGGGGTTTCCATTTTCAAACGGGAATC
CTAGCAGCTTCCCTGCTCCCACCTGCGGAAGAGGAGACACTTTTGTACATTCTGACTTTCTGCAGACAGG
TAAAAAGGCGTACACAGACCTTTGGGGATGAAAGAGAAACTAAAACTGAACCAAGTGTACAGGTTACAGC
GTCACAGTCAAGGCACCTAACAGACCCTACTTACTGTTTGTTCTTAAACATGAATGACTGCAGATCCCTG
CCAGAGACGGTAAGTGAAAAAACAGCTACCTCCTATTTTCTTTACGTATTTCCCATCAAACGCCTGTCTT
GGGTTGGGTTAAGAATTACAGAAGGTAAACGGAGTCAATTCCAAGTCACTGTTACGCCTTCTATGGCAATCA
AACACGCC

>DEX0367_009.nt.1
ATAAGTGTCTTAACTTGGGCTGTCTTCACTCCACCCCTCCCGAGTAGGTACTTCAGTTGTGCCCATTCTA
CAGACAGGGAAGCTGAGGCTGGAGAGGTCAGGACTCGCCTGAGGTCCTACGGTCTCCCCTGGGACCTGGC
TGAGGATGGGGGCCGGGCTGGGCCCTCAGGGTTGGAGACTCTGACCCCCTACTCCCGACCCCCAGCTTC
ACGTGGAGTGACGCACGCCTGCACCGAGGCCTGGTGACCCTGCTGACTGGTGAAATTGTTGACGCCTTCA
GCCTTGAGTTCCGGACGCTGTACGCGGCCTCCTGCCCGCTCCCACCTGCGCCCCCCCAGAAACCCTCGGT
CATAGGTGGCCTGCAGCGGGGCCGCAGCCTGCACCGCGTGTCCCTGCCGCGTCTCCCCTGCTACCCCGGGGC
CGGCACTCAGTGACATTCTAAGGAGTGTGCAGCGCGCCCTGCCGCGTCTCCCCTGCTACCCCGGGGC
CCGCTCCATGTGGGACCTAAGCCGCCTGTCCCAGCTGTCTGGCTCCAGTGATGGGGACAACGAGCTCAAG
AAATCCTGGGGCTCCAAGGACACTCCAGCCAAGGCTCTGATGAGGCAGCGGGGCACTGGAGGAGGCCCCT
GGGGGGAAGTGGACTCCCGACCTCCGTGGGGCGGTGCCCTGCCCCTGCCCCCCGCCCACCGCCTCCGCTA
TCTGTCCCCAGCCCGAAGGCGGTTCGGTGGGGATGCTACATTCAAACTTCAAGAGCCCAGAGGCGTCAGG
CCGTCAGACTGGGCCCCCCGGGCAGGACTTGGGGGGCAACCCTGAACAGGAGCCCAAGCCAAATCTGCTG
GCCCACCCAAGGAGCAGTGCCGGACAACTGTCGGCTCAGAGCCCTCTGCCTCAATGCCCTGCCTTGGACA
AGGCTTCTGGCCTCGGCGTCTGTCAAGCAGAAGGGATGTTTGTCACTTGTTCTGCAACTTGGGCCGGGAG
TGGTGGCTCATGCCTGTAATCCCTGCACTTTGGGAGGCTAAGGTGGGAGGGTTGCTTAAGGCCAGGAGTT
GGAGACCAGCCTGGGCAACATAGTGAGACCCCATCTCTACAATTAAAAAAAACAAATTAGTTGGGTGTGGT
GGCGCACGCCTGTGGTCCCAGCTACGCAGGAGGCTGAGGTGGGAGGATCCCTCGAGCTCAGGAGGTGGAG
GCAGCAGTGAGCCAAGATCATGCCCTTGCACTCCAGCCTGGGCAACACAGTGAGAGGCCATCTCTTTA

>DEX0367_010.nt.1
CCCTTTTCAATGAAAAGTCACTGAGCTTGGTGAGTCAAGATTCACTGAAATATTCTTCCATGTTGATGGG
CCTCTTTTGGTGAGGTGACCTTATTCCTGTCCTGCCCCCTAAGGGTGGGCAGGGTGGACAGAACGCAGCCT
GGACTTCCGGACGCAGCCCCCTACCGCCCCGGACTCGCCTACCCCTCTGGCCTCTGAGGAGGGCCCTAAG
CCCAGGACAGACGTGGAAAAAGGGCAGTAAAAAGGATGCAGTTCTTCTCTTTAATCCCTAACCCAATTTC
TCGGGAAAAGCTGCGTTCCCACACAGAAGGCGAACCCTGTAAAGCTGTGGCTTGGCTGCCGCGCACGCAG
TCGCTTTACGAATTCCCAAAGCATCTAGCGGAAGTTTAGACCGTTCACTAGGGGTCACATGGGCCACTTC
GCGAGCAGCTGGCGAACAGTTCCTTTCAGTTTGAAAGTATCTTTGAAACTCGGAAATTGTGCTCCTTGGT
CTTTCAGTTCATAAGCCTAGAGAGTGGCTACTGAAATTGGAATCGTCATCTTCTTCCCTCGGCCCCAACT
AGTAGACCACGACCAGAAGACACATCCTAACCTACACCTACCGTAGTGTCACACCCGGCACAGCGCGACC
AGGTCCACCAACTGCGCCTGCGCGAACCGGCCGCTCCCGCGACATCACAGGAAGGCTGAGAGAAAGGGCA
GGTCCCCGCTGCGTGGGGCTGGGCTTAAAGGCTCGTCGCGCCTGCGCAAAGGACCTGACGACGTGCTGCG
TCGTTACTTTTGAAACGCTTGGCGGGGAAGTGCTGTTGGAGCCGCTGTGGTTGCTGTCCGCGGAGTGGAA
GCGCGTGCTTTTTGTTTGTGTCCCTGGCCATGGCGCTGCAGCTCTCCCGGGGAGCAGGGAATCACCCTGCGC
GGGAGCGCCGAAATCGTGGCCGAGTTCTTCTCATTCGGCATCAACAGCATTTTATATCAGCGTGGCATAT
ATCCATCTGAAACCTTTACTCGAGTGCAGAAATACGGACTCACCTTGCTTGTAACTACTGATCTTGAGCT
CATAAAATACCTAAATAATGTGGTGGAACAACTGAAAGTGCACCCAGAGAAAAGTCTCAGAAAGCTATCC

AGGATGAAATCCGTTCAGTGATCAGACAGATCACAGCTACGGTGACATTTCTGCCACTGTTGGAAGTTTC
TTGTTCATTTGATCTGCTGATTTATACAGACAAAGATTTGGTTGTACCTGAAAAATGGGAAGAGTCGGGA
CCACAGTTTATTACCAATTCTGAGGAAGTCCGCCTTCGTTCATTTACTACTACAATCCACAAAGTAAATA
GCATGGTGGCCTACAAAATTCCTGTCAATGACTGAGGATGACATGAGGAAATAATGTAATTGTAATTTT

5   GAAATGTGGTTTTCCTGAAATCAGGTCATCTATAGTTGATATGTTTTATTTCATTGGTTAATTTTTACAT
GGAGAAAACCAAAATGATACTTACTGAACTGTGTGTAATTGTTCCTTTTATTTTTTTGGTACCTATTTGA
CTTACCATGGAGTTAACATCATGAATTTATTGCACATTGTTCAAAAGGAACCAGGAGGTTTTTTGTCAA
CATTGTGATGTATATTCCTTTGAAGATAGTAACTGTAGATGGGAAAAACTTGTGCTATAAAGCTAGATGCT
TTCCTAAATCAGATGTTTTGGTCAAGTAGTTTGACTCAGTATAGGTAGGGAGATATTTAAGTATAAAATA

10 CAACAAAGGAAGTCTAAATATTCAGAATCTTTGTTAAGGTCCTGAAAGTAACTCATAATCTATAAACAAT
GAAATATTGCTGTATAGCTCCTTTTGACCTTCATTTCATGTATAGTTTTCCCTATTGAATCAGTTTCCAA
TTATTTGACTTTAATTTATGTAACTTGAACCTATGAAGCAATGGATATTTGTACTGTTTAATGTTCTGTG
ATACAGAACTCTTAAAAATGTTTTTTCATGTGTTTTATAAAATCAAGTTTTAAGTGAAAGTGAGGAAATA
AAGTTAAGTTTGTTTTAAATTTGTCTTAAAA ·

15
>DEX0367_011.nt.1
TCGAAYMGTAACCGTTCGTACGAGAATCGCTGTCCTCTCCTTGCGGCCGCAGTGTGGTGGAATTCTTTTC
CGACACGTAGTAGTACTGACGCTTGGGAACGGCCTGGGCCAGAATCTCGATCTGGCGAGCGTTCAGGCCC
ATGCGGCGGTACAGGGCCGCCGTGTCCTCATCCCTGGCGTAAATATTCGGCAGGAAAATCTTGGTCGCGG

20 TCGATTCCACGATCACGTCCAGGATGCCGCTGTTGGCGGCGTCGGACAGGCTCTGCGTTGCCATCAGCAC
AAGGCAGTTGGCCTTACGCAGCACCTTGAGCCATTCCCTGATCTTCGCGCGGAATGCCGGGTGGCCGAGC
ATCAACCAGGCTTCGTCCAGGATGATGACGGCCGCCAGTGTGGTGGAATTCTAACCCTGAAAAAGATATA
TTTGTGGTGCGGGAAAATGGGACGACGTGTCTCATGGCAGAGTTTGCAGCCAAATTTATTGTACCTTATG
ATGTGTGGGCCAGCAACTACGTAGATCTGATCACAGAACAGGCCGATATCGCATTGACCCGGGGAGCTGA

25 GGTGAAGGGCCGCTGTGGCCACAGCCAGTCGGAGCTGCAAGTGTTCTGGGTGGATCGCGCATATGCACTC
AAAATGCTCTTTGTAAAGGAAAGCCACAACATGTCCAAGGGACCTGAGGCGACTTGGAGGCTGAGCAAAG
TGCAGTTTGTCTACGACTCCTCGGAGAAAACCCACTTCAAAGACGCAGTCAGTGCTGGGAAGCACACAGC
CAACTCGCACCACCTCTCTGCCTTGGTCACCCCGCTGGGAAGTCCTATGAGTGTCAAGCTCAACAAACC
ATTTCACTGGCCTCTAGTGATCCGCAGAAGACGGTCACCATGATCCTGTCTGCGGTCCACATCCAACCTT

30 TTGACATTATCTCAGATTTTGTCTTCAGTGAAGAGCATAAATGCCCAGTGGATGAGCGGGAGCAACTGGA
AGAAACCTTGCCCCTGATTTTGGGGCTCATCTTGGGCGTCGTCATCATGGTAACACTCGCGATTTACCAC
GTCCACCACAAAATGACTGCCAACCAGGTGCAGATCCCTCGGGACAGATCCCAGTATAAGCACATGGGCT
AGAGGCCGTTAGGCAGGCACCCCCTATTCCTGCTCCCCCAACTGGATCAGGTAGAACAACAAAAGCACTT
TTCCATCTTGTACACGAGATACACCAACATAGCTACAATCAAACAGGCCTGGGTATCTGAGGCTTGCTTG

35 GCTTGTGTGTCCATGCTTAAACCCACGGAAGGGGGAGACTCTTTCGGATTTGTAGGGTGAAATGGCAATTAT
TCTCTCCATGCTGGGGAGGAGGGGAGGAGGGTCTCAGCACAGCTTTCGTGCTCATGGTGGCTTGGCTTTGA
CTCTCCAAAGAGCAATAAATGCCACTTGGAGCTGTATCTGGCCCCAAAGTTTAGGGATTGAAAACATGCT
TCTTTGAGGAGGAAACCCCTTTAGGTTCAGAAGAATATGGGGTGCTTTGCTCCCTTGGACACAGCTGGCT
TATCCTATACAGTTGTCAATGCACACAGAATACAACCTCATGCTCCCTGCAGCAAGACCCCTGAAAGTGA

40 TTCATGCTTCTGGCTGGCATTCTGCATGTTTAGTGATTGTCTTGGGAATGTTTCACTGCTACCCGCATCC
AGCGACTGCAGCACCAGAAAACGACTAATGTAACTATGCAGAGTTGTTTGGACTTCTTCCTGTGCCAGGT
CCAAGTCGGGGGACCTGAAGAATCAATCTGTGTGAGTCTGTTTTTCAAAATGAAATAAAACACACTATTC
TCTGGC

45 >DEX0367_012.nt.1
ATAATTTGGAATGTGGGTGGAGAGGGGATCATAGTTCTCCCTGAGTGAGACTCACCTGCTCCTCTGGCCC
CTGGTCCTGTCCTGTTCTCCAGCATGGTGTGTCTGAAGCTCCCTGGAGGCTCCTGCATGGCAGCTCTGAC
AGTGACACTGATGGTGCTGAGCTCCCACTGGCTTTGGCTGGGGACACCCAACCACGTTTCTTGTGGCAG
GATAAGTATGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGAAAGAGTCATCTATAACC

50 AAGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGGAGCTAGGGCGGCCTGT
CGCCGATCCATCCTGAGGTGACTGTGTATCCTGCCAAGACTCAGCCCCTGCAGCACCACAACCTCCTGGT
CTGCTCTGTGAGTGGTTTCTATCCAGGCAGCATTGAAGTCAGGTGGTTCCGGAACGGCCAGGAAGAGAAG
GCTGGGGTGGTGTCCACAGGCCTGATCCAGAATGGAGACTGGACCCAAGTGTGATGAGCCCTCTCACAGT

55 GGAATGGAGAGCACGGTCTGAATCTGCACAGAGCAAGATGCTGAGTGGAGTCGGGGGCTTTGTGCTGGGC
CTGCTCTTCCTTGGGGCCGGGTTGTTCATCTACTTCAGGAATCAGAAAGGACACTCTGGACTTCAGCCAA
CAGGATTCCTGAGCTGAAGTGAAGATGACCACATTCAAGGAAAAACCTTCTGCCCCAGCTTTGCAGGATG
AAACACTTCCCCGCTTGGCTCTCATTCTTCCACAAGAGAGACCTTTCTCCGGACCTGGTTGCTACTGGTT
CAGCAACTCTGCAGAAAATGTCCTCCCCTGTGGCTGCCTCAGCTCATGCCTTTGGCCTGAAGTCCCAGCA

60 TTGATGGCAGCCCCTCATCTTCCAAGTTTTGTGCTCCCCTTTACCTAACGCTTCCTGCCTCCCATGCATC

TGTACTCCTCCTGTGCCACAAACACATTACATTATTAAATGTTTCTCAAACATGGA

>DEX0367_012.nt.2
ATAATTTGGAATGTGGGTGGAGAGGGGATCATAGTTCTCCCTGAGTGAGACTCACCTGCTCCTCTGGCCC
CTGGTCCTGTCCTGTTCTCCAGCATGGTGTGTCTGAAGCTCCCTGGAGGCTCCTGCATGGCAGCTCTGAC
AGTGACACTGATGGTGCTGAGCTCCCCACTGGCTTTGGCTGGGGACACCCAACTCCATCCTGAGGTGACT
GTGTATCCTGCCAAGACTCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGTGAGTGGTTTCTATC
CAGGCAGCATTGAAGTCAGGTGGTTCCGGAACGGCCAGGAAGAGAAGGCTGGGGGTGGTGTCCACAGGCCT
GATCCAGAATGGAGACTGGACCTTCCAGACCCTGGTGATGCTGGAAACAGTTCCTCGGAGTGGAGAAGTT
TACACCTGCCAAGTGGAGCACCCAAGTGTGATGAGCCCTCTCACAGTGGAATGGAGAGCACGGTCTGAAT
CTGCACAGAGCAAGATGCTGAGTGGAGTCGGGGGGCTTTGTGCTGGGCCTGCTCTTCCTTGGGGCCGGGTT
GTTCATCTACTTCAGGAATCAGAAAGGACACTCTGGACTTCAGCCAACAGGATTCCTGAGCTGAAGTGAA
GATGACCACATTCAAGGAAAAACCTTCTGCCCCAGCTTTGCAGGATGAAACACTTCCCCGCTTGGCTCTC
ATTCTTCCACAAGAGAGACCTTTCTCCGGACCTGGTTGCTACTGGTTCAGCAACTCTGCAGAAAATGTCC
TCCCCTGTGGCTGCCTCAGCTCATGCCTTTGGCCTGAAGTCCCAGCATTGATGGCAGCCCCTCATCTTCC
AAGTTTTGTGCTCCCCTTTACCTAACGCTTCCTGCCTCCCATGCATCTGTACTCCTCCTGTGCCACAAAC
ACATTACATTATTAAATGTTTCTCAAACATGGA

>DEX0367_013.nt.1
CATGATTACGCACAAGCTTCAATATTCACTTTTATTATTTATAAAAAAAAATTAAGGTACAGACCCCTAG
GCGTCCTGACAGGTGGATTTCGACAAGGTCATTGTGCCCTGCCAAGGCACAGCGTAGATCTGGAAAGAGC
AGAATGCCTCCGGCCCCCAGGGACCCCCGATGATGCCACCTTACCCTGGTCCCGGACATTGCCATCAACAG
CTGAACAGCCTGCCCACCATGGAAGGACCCCCAACCTTCAACCCGCCTGTGCCATATTTCGGGAGGCTGC
AAGGAGGGCTCACAGCTCGAAGAACCATCATCATCAAGGGCTATGTGCCTCCCACAGGCAAGAGCTTTGC
TATCAACTTCAAGGTGGGCTCCTCAGGGGACATAGCTCTGCACATTAATCCCCGCATGGGCAACGGTACC
GTGGTCCGGAACAGCCTTCTGAATGGCTCGTGGGGATCCGAGGAGAAGAAGATCACCCACAACCCATTTG
GTCCCGGACAGTTCTTTGATCTGTCCATTCGCTGTGGCTTGGATCGCTTCAAGGTTTACGCCAATGGCCA
GCACCTCTTTGACTTTGCCCATCGCCTCTCGGCCTTCCAGAGGGTGGACACATTGGAAATCCAGGGTGAT
GTCACCTTGTCCTATGTCCAGATCTAATCTATTCCTGGGGCCATAACTCATGGGAAAACAGAATTATCCC
CTAGGACTCCTTTCTAAGCCCCTAATAAAATGTCTGAGGGTGTCTCATGAAAAAAAA

>DEX0367_013.nt.2
GGATTCAGAGGTCCAGCCCACTGGCCTCTTCCCACTGGACTCCATCTGACTCCCCCATCCCTCTCTGTCC
CCAGCTTTGCTATCAACTTCAAGGTGGGCTCCTCAGGGGACATAGCTCTGCACATTAATCCCCGCATGGG
CAACGGTACCGTGGTCCGGAACAGCCTTCTGAATGGCTCGTGGGGATCCGAGGAGAAGAAGATCACCCAC
AACCCATTTGGTCCCGGACAGTTCTTTGATCTGTCCATTCGCTGTGGCTTGGATCGCTTCAAGGTTTACG
CCAATGGCCAGCACCTCTTTGACTTTGCCCATCGCCTCTCGGCCTTCCAGAGGGTGGACACATTGGAAAT
CCAGGGTGATGTCACCTTGTCCTATGTCCAGATCTAATCTATTCCTGGGGCCATAACTCATGGGAAAACA
GAATTATCCCCTAGGACTCCTTTCTAAGCCCCTAATAAAATGTCTGAGGGTGTCTCATGAAAAAAAA

>DEX0367_014.nt.1
CATCCAATGAGACACCAGCTCAGGTTACTGCAGGGGACACACTATAAAGCCCTGAGCTCAGGGAGGAGCT
CCCTCCAGGCTCTATATTTAGAGCGGGTAGGGGAGCGCAGCGGCCAGATACCTCAGCGCTACCTGGCGG
AACTGGATTTCTCTCCCGCCTGCCGGCCTGCCTGCCACAGCCGGACTCCGCCACTCCGGTAGCCTCATGG
CTGCAACCTGTGAGATTAGCAACATTTTTAGCAACTACTTCAGTGCGATGTACAGCTCGGAGGACTCCAC
CCTGGCCTCTGTTCCCCCTGCTGCCACCTTTGGGGCCGATGACTTGGTACTGACCCTGAGCAACCCCCAG
ATGTCATTGGAGGGTACAGAGAAGGCCAGCTGGTTGGGGGAACAGCCCCAGTTCTGGTCGAAGACGCAGG
TTCTGGACTGGATCAGCTACCAAGTGGAGAAGAACAAGTACGACGCAAGCGCCATTGACTTCTCACGATG
TGACATGGATGGCGCCACCCTCTGCAATTGTGCCCTTGAGGAGCTGCGTCTGGTCTTTGGGCCTCTGGGG
GACCAACTCCATGCCCAGCTGCGAGACCTCACTTCCAGCTCTTCTGATGAGCTCAGTTGGATCATTGAGC
TGCTGGAGAAGGATGGCATGGCCTTCCAGGAGGCCCTAGACCCAGGGCCCTTTGACCAGGGCAGCCCCTT
TGCCCAGGAGCTGCTGGACGACGGTCAGCAAGCCAGCCCCTACCACCCCGGCAGCTGTGGCGCAGGAGCC
CCCTCCCCTGGCAGCTCTGACGTCTCCACCGCAGGGACTGGTGCTTCTCGGAGCTCCCACTCCTCAGACT
CCGGTGGAAGTGACGTGGACCTGGATCCCACTGATGGCAAGCTCTTCCCCAGCGGTGAGTCGAGGGAGGT
CCCCAAGAGGGCGTCCCATTTAGCAATGCACAGGGGGCCCGGCTCTTCCTGCAGCCTTTTCCTGTAGAGG
GGCTACTCTCCCTAACTCCCCTCTTGCCCCTCCTTGACCTTCCACCACCGTCCCCACAGATGGTTTTCGT
GACTGCAAGAAGGGGGATCCCAAGCACGGGAAGCGGAAACGAGGCCGGCCCCGAAAGCTGAGCAAAGAGT
ACTGGGACTGTCTCGAGGGCAAGAAGAGCAAGCACGGTGAGCTCCGGGGGCACGTGGGTCCTCCCTGCGC
CGGGCTGAGCGGCTTCCTGGGGCACTGCGGGTTGTTGCAGGTATCCCTTCTCCCGTTTTTCTCTGGCCTCC
GCATGGCCTTTGGTAAGGCTGTGCACAAGCTGGGGGCTCTATGGTATCGGTCACCACCTAATTGCAGAGC

CTGGCTTGGTGGTCCTGGAGAGGAGGAGGAAATAAGGCTCCCAGTGGGAGGCTCATGGTACCAGAGTCCT
GTCCACTGACTCCAGTGTCCTGTCCACTGACTCCAGTTCTCTCTGCACTTGGCCACTGTCCTGCCCTCTG
GGACACCCTCAATGTGAGGAGGCAGCTGGTGGGTCTTAGGTGGGCTGAGGAGAAAAGCAGTCACTGCAGT
ACCCGCACAGAGGGCACTGCGGGGTCTCTGGAGAGGCTTGCTGCATGCTGTGGCCAAGTCAGCAGTGCAC

5 TGGGGCGGGCAGGGCTGGCTGGCCTTGGGTGAGAGGGGACACCTGGATGGCAAACTGATGGAGGCTGGCC
TTGCAGCGCCCAGAGGCACCCACCTGTGGGAGTTCATCCGGGACATCCTCATCCACCGGGAGCTCAACGA
GGGCCTCATGAAGTGGGAGAATCGGCATGAAGGCGTCTTCAAGTTCCTGCGCTCCGAGGCTGTGGCCCAA
CTATGGGGCCAAAAGAAAAAGAACAGCAACATGACCTACGAGAAGCTGAGCCGGGCCATGAGGTACTACT
ACAAACGGGAGATCCTGGAACGGGTGGATGGCCGGCGACTCGTCTACAAGTTTGGCAAAAACTCAAGCGG

10 CTGGAAGGAGGAAGAGGTTCTCCAGAGTCGGAACTGAGGGTTGGAACTATACCCGGGACCAAACTCACGG
ACCACTCGAGGCCTGCAAACCTTCCTGGGAGGACAGGCAGGCCGACATGGCCCCTCCACTGGGGAATGCTC
CCAGCTGTGCTGTGGAGAGAAGCTGATGTTTTGGTGTATTGTCAGCCATCGTCCTGGGACTCGGAGACTA
TGGCCTCGCCTCCCCACCCTCCTCTTGGAATTACAAGCCCTGGGGTTTGAAGCTGACTTTATAGCTGCAA
GTGTATCTCCTTTTTATCTGGTGCCTCCTCAAACCCAGTCTCAGACACTAAATGCAGACAACACCTTCCTC

15 CTGCAGACACCTGGACTGAGCCAAGGAGGCCTGGGGAGGCCCTAGGGGAGCACCGTGATGGAGAGGACAG
AGCAGGGGCTCCAGCACCTTCTTTCTGGACTGGCGTTCACCTCCCTGCTCAGTGCTTGGGCTCCACGGGC
AGGGGTCAGAGCACTCCCTAATTTATGTGCTATATAAATATGTCAGATGTACATAGAGATCTATTTTTTC
TAAAACATTCCCCTCCCCACTCCTCTCCCACAGAGTGCTGGACTGTTCCAGGCCCTCCAGTGGGCTGATG
CTGGGACCCTTAGGATGGGGCTCCCAGCTCCTTTCTCCTGTGAATGGAGGCAGAGACCTCCAATAAAGTG

20 CCTTCTGGGCTTTTTCTAACCTTTGTCTTAGCTACCTGTGTACTGAAATTTGGGCCTTTGGATCGAATAT
GGTCAAGAGGTTGGAGGGGAGGAAAATGAAGGTCTACCAGGCTGAGGGTGAGGGCAAAGGCTGACGAAGA
GGGGAGTTACAGATTTCCTGTAGCAGGTGTGGGCTTACAGACACATGGACTGGGCTGGGAGGCGAGCAAA
GGAAGCAGCTGAGACTGTTGGAGAACGCTTACAAGACTTCATGCAAGCAAGGACATGAACTCAGAACACT
GAGGTCAGAAGCATCCTGCTGTCATGACACCGCTCGAGTGACCTTGACCTTGACCAAGTCTGTCCTGTTT

25 AGGACTGATTTTTCCTATTAGGCTAGGGTTTGGACCTGATGTTCTCAAGATGTCTAGAATTGCATGGCTG
GCCTTGTGGAATAGATGGTTTTGCATTCCAGCCAAGTGTGCTGTAAACTGTATATCTGTAATATGAATCC
CAGCTTTTGAGTCTGACAAAATCAGAGTTAGGATCTTGTAAAGGAAAAAAAAAAAAAAACAAAACAAAATG
GAGATGAGTACTTGCTGAGAAAGAATGAGGGAAGGAGTTGGCATTTGTTGAAAGTATAGTCTTTTTCTCT
TTTTTTTTTAATTGCAACTTTTACTTTAGATTTAGGAGGTCGTGCGCAGGTTTGTTACATGGGTATATTG

30 TGTGATGCTGAGCTTGGGATGCGAATGATCCTGTCACCCAGGTAGTGAGTATAGCACCCAGTGAAACTGT
AGTCTCATGCCAGGCACTGTGCTAGCCCACTCTGGCTCATTTAATCCTCTCCTAAGAAGAGAGGAGACAC
AGCGTCCCCATTTGACAGATGCAGAAAGAGGTTCCACAGGTGTGCCTTGATTCTGTCCTAAAACCGTTTC
CCGGAAGCTTTTCCTGGTGTGGGCGCTTCTAACCTAATCCTCAATCGATTCCAGAACTATTACTCTGTTT
CCACAGTGATACTGTGTCTAGGTTTTAGGGAGGACAGTTCATTGATGTTACTTAAGAATGCTTTCCAGGT

35 GGAAAGTTCCTTAAGTTTGAGGCTTCAAATTCCATACAGCACATTAAAATCCCATTCATGAGTTTGAAAT
ACTGCTCTGTTGTCTTGGAAATACCAATCAGATTGTTGGCTGAAGTGATGTGGATAAAGAAGGGATCTTA
GAAAAACTAAAGCTTGTGTG

>DEX0367_014.nt.2
40 CATCCAATGAGACACCAGCTCAGGTTACTGCAGGGGACACACTATAAAGCCCTGAGCTCAGGGAGGAGCT
CCCTCCAGGCTCTATATTTAGAGCCGGGTAGGGGAGCGCAGCGGCCAGATACCTCAGCGCTACCTGGCGG
AACTGGATTTCTCTCCCGCCTGCCGGCCTGCCTGCCACAGCCGGACTCCGCCACTCCGGTAGCCTCATGG
CTGCAACCTGTGAGATTAGCAACATTTTTAGCAACTACTTCAGTGCGATGTACAGCTCGGAGGACTCCAC
CCTGGCCTCTGTTCCCCCTGCTGCCACCTTTGGGGCCGATGACTTGGTACTGACCCTGAGCAACCCCCAG

45 ATGTCATTGGAGGGTACAGAGAAGGCCAGCTGGTTGGGGGAACAGCCCCAGTTCTGGTCGAAGACGCAGG
TTCTGGACTGGATCAGCTACCAAGTGGAGAAGAACAAGTACGACGCAAGCGCCATTGACTTCTCACGATG
TGACATGGATGGCGCCCACCCTCTGCAATTGTGCCCTTGAGGAGCTGCGTCTGGTCTTTGGGCCTCTGGGG
GACCAACTCCATGCCCAGCTGCGAGACCTCACTTCCAGCTCTTCTGATGAGCTCAGTGGATCATTGAGC
TGCTGGAGAAGGATGGCATGGCCTTCCAGGAGGCCCTAGACCCAGGGCCCTTTGACCAGGGCAGCCCCTT

50 TGCCCAGGAGCTGCTGGACGACGGTCAGCAAGCCAGCCCCTACCACCCCGGCAGCTGTGGCGCAGGAGCC
CCCTCCCCTGGCAGCTCTGACGTCTCCACCGCAGGAACAGGAACAGGCTGGGAAGTGTGTCCTGAGAGCC
AGCAGCGTGGTTGAGCAGAGGGTGGGCCGGCAGGGGACTTACTCTGACCCCGCCCCCCGGACTGGTGCTT
CTCGGAGCTCCCACTCCTCAGACTCCGGTGGAAGTGACGTGGACCTGGATCCCACTGATGGCAAGCTCTT
CCCCAGCGATGGTTTTCGTGACTGCAAGAAGGGGGATCCCAAGCACGGGAAGCGGAAACGAGGCCGGCCC

55 CGAAAGCTGAGCAAAGAGTACTGGGACTGTCTCGAGGGCAAGAAGAGCAAGCACGCGCGCCCAGAGGCACCC
ACCTGTGGGAGTTCATCCGGGACATCCTCATCCACCCGGGAGCTCAACGAGGGCCTCATGAAGTGGGAGAA
TCGGCATGAAGGCGTCTTCAAGTTCCTGCGCTCCGAGGCTGTGGCCCAACTATGGGGCCAAAAGAAAAAG
AACAGCAACATGACCTACGAGAAGCTGAGCCGGGCCATGAGGTACTACTACAAACGGGAGATCCTGGAAC
GGGTGGATGGCCGGCGACTCGTCTACAAGTTTGGCAAAAACTCAAGCGGCTGGAAGGAGGAAGAGGTTCT

60 CCAGAGTCGGAACTGAGGGTTGGAACTATACCCGGGACCAAACTCACGGACCACTCGAGGCCTGCAAACC

```
     TTCCTGGGAGGACAGGCAGGCCAGATGGCCCCTCCACTGGGGAATGCTCCCAGCTGTGCTGTGGAGAGAA
     GCTGATGTTTTGGTGTATTGTCAGCCATCGTCCTGGGACTCGGAGACTATGGCCTCGCCTCCCCACCCTC
     CTCTTGGAATTACAAGCCCTGGGGTTTGAAGCTGACTTTATAGCTGCAAGTGTATCTCCTTTTATCTGGT
     GCCTCCTCAAACCCAGTCTCAGACACTAAATGCAGACAACACCTTCCTCCTGCAGACACCTGGACTGAGC
5    CAAGGAGGCCTGGGGAGGCCCTAGGGGAGCACCGTGATGGAGAGGACAGAGCAGGGGCTCCAGCACCTTC
     TTTCTGGACTGGCGTTCACCTCCCTGCTCAGTGCTTGGGCTCCACGGGCAGGGGTCAGAGCACTCCCTAA
     TTTATGTGCTATATAAATATGTCAGATGTACATAGAGATCTATTTTTTCTAAAACATTCCCCTCCCCACT
     CCTCTCCCACAGAGTGCTGGACTGTTCCAGGCCCTCCAGTGGGCTGATGCTGGGACCCTTAGGATGGGGC
     TCCCAGCTCCTTTCTCCTGTGAATGGAGGCAGAGACCTCCAATAAAGTGCCTTCTGGGCTTTTTCTAACC
10   TTTGTCTTAGCTACCTGTGTACTGAAATTTGGGCCTTTGGATCGAATATGGTCAAGAGGTTGGAGGGGAG
     GAAAATGAAGGTCTACCAGGCTGAGGGTGAGGGCAAAGGCTGACGAAGAGGGGAGTTACAGATTTCCTGT
     AGCAGGTGTGGGCTTACAGACACACATGGACTGGGCTGGGAGGCGAGCAAAGGAAGCAGCTGAGACTGTTGG
     AGAACGCTTACAAGACTTCATGCAAGCAAGGACATGAACTCAGAACACTGAGGTCAGAAGCATCCTGCTG
15   TCATGACACCGCTCGAGTGACCTTGACCTTGACCAAGTCTGTCCTGTTTAGGACTGATTTTTCCTATTAG
     GCTAGGGTTTGGACCTGATGTTCTCAAGATGTCTAGAATTGCATGGCTGGCCTTGTGGAATAGATGGTTT
     TGCATTCCAGCCAAGTGTGCTGTAAACTGTATATCTGTAATATGAATCCCAGCTTTTGAGTCTGACAAAA
     TCAGAGTTAGGATCTTGTAAAGGAAAAAAAAAAAAAAACAAAACAAAATGGAGATGAGTACTTGCTGAGAA
     AGAATGAGGGAAGGAGTTGGCATTTGTTGAAAGTATAGTCTTTTTCTCTTTTTTTTTAATTGCAACTTT
20   TACTTTAGATTTAGGAGGTCGTGCGCAGGTTTGTTACATGGGTATATTGTGTGATGCTGAGCTTGGGATG
     CGAATGATCCTGTCACCCAGGTAGTGAGTATAGCACCCAGTGAAACTGTAGTCTCATGCCAGGCACTGTG
     CTAGCCCACTCTGGCTCATTTAATCCTCTCCTAAGAAGAGGAGGAGACACAGCGTCCCCATTTGACAGATG
     CAGAAAGAGGTTCCACAGGTGTGCCTTGATTCTGTCCTAAAACCGTTTCCCGGAAGCTTTTCCTGGTGTG
     GGCGCTTCTAACCTAATCCTCAATCGATTCCAGAACTATTACTCTGTTTCCACAGTGATACTGTGTCTAG
     GTTTTAGGGAGGACAGTTCATTGATGTTACTTAAGAATGCTTTCCAGGTGGAAAGTTCCTTAAGTTTGAG
25   GCTTCAAATTCCATACAGCACATTAAAATCCCATTCATGAGTTTGAAATACTGCTCTGTTGTCTTGGAAA
     TACCAATCAGATTGTTGGCTGAAGTGATGTGGATAAAGAAGGGATCTTAGAAAAACTAAAGCTTGTGTG
```

```
>DEX0367_015.nt.1
     AATTCAAGAATGTCACTGCCAGTTAAACCTGAATTGTTAGGGGATTTAGAAATACCAGCAGTGCCCATAC
30   TCCATAGCATGGTGCAAAAATTCCCAGGCGTGTCATTTGGGATCAGCACTGATTCTGAGGTTCTGACACA
     CTACAACATCACTGGGAACACCATCTGCCTCTTTCGCCTGGTAGACAATGAACAACTGAATTTAGAGGAC
     GAAGACATTGAAAGCATTGATGCCACCAAATTGAGCCGTTTCATTGAGATCAACAGCCTCCACATGGTGA
     CAGAGTACAACCCTGTGACTGTGATTGGGTTATTCAACAGCGTAATTCAGATTCATCTCCTCCTGATAAT
     GAACAAGGCCTCCCCAGAGTATGAAGAGAACATGCACAGATACCAGAAGGCAGCCAAGCTCTTCCAGGGG
35   AAGATTCTCTTTATTCTGGTGGACAGTGGTATGAAAGAAAATGGGAAGGTGATATCATTTTTCAAACTAA
     AGGAGTCTCAACTGCCAGCTTTGGCAATTTACCAGACTCTAGATGACGAGTGGGATACACTGCCCACAGC
     AGAAGTTTCCGTAGAGCATGTGCAAAACTTTTGTGATGGATTCCTAAGTGGAAAATTGTTGAAAGAAAAT
     CGTGAATCAGAAGGAAAGACTCCAAAGGTGGAACTCTGACTTCTCCTTGGAACTACATATGGCCAAGTAT
     CTACTTTATGCAAAGTAAAAAGGCACAACTCAAATCTCAGAGACACTAAACAACAGGATCACTAGGCCTG
40   CCAACCACACACACACGCACGTGCACACGCACGCACGCGTGCACACACACACGCGCACACACACACAC
     ACACACACAGAGCTTCATTTCCTGTCTTAAAATCTCGTTTTCTCTTCTTCCTTCTTTTTAAATTTCATATC
     CTCACTCCCTATCCAATTTCCTTCTTATCGTGCATTCATACTCTGTAAGCCCATCTGTAACACACCTAGA
     TCAAGGCTTTAAGAGACTCACTGTGATGCCTCTATGAAAGAGAGGCATTCCTAGAGAAAGATTGTTCCAA
     TTTGTCATTTAATATCAAGTTTGTATACTGCACATGACTTACACACAACATAGTTCCTGCTCTTTTAAGG
45   TTACCTAAGGGTTGAAACTCTACCTTCTTTCATAAGCACATGTCCGTCTCTGACTCAGGATCAAAAACCA
     AAGGATGGTTTTAAACACCTTTGTGAAATTGTCTTTTTGCCAGAAGTTAAAGGCTGTCTCCAAGTCCCTG
     AACTCAGCAGAAATAGACCATGTGAAAACTCCATGCTTGGTTAGCATCTCCAACTCCCTATGTAAATCAA
     CAACCTGCATAATAAATAAAAGGCAATCATGTTATAGGAGAAGGCTGGCATCTCTTTGCA
```

```
50   >DEX0367_015.nt.2
     ACCAGACCTAGGAATTCCTAGAAAAATAGTCAGGAAGCATTTAGACACATCAAATGTTAAACGAGTCCTG
     ATTATGATGATAATGATGATGATTTTGGTGGTTGCAATAGCAAAGCCTTAAGTATGAAGGAGACGTGCCA
     GCTGGAAATACAGGTAGACAATGAACAACTGAATTTAGAGGACGAAGACATTGAAAGCATTGATGCCACC
     AAATTGAGCCGTTTCATTGAGATCAACAGCCTCCACATGGTGACAGAGTACAACCCTGTGACTGTGATTG
55   GGTTATTCAACAGCGTAATTCAGATTCATCTCCTCCTGATAATGAACAAGGCCTCCCCAGAGTATGAAGA
     GAACATGCACAGATACCAGAAGGCAGCCAAGCTCTTCCAGGGGAAGATTCTCTTTATTCTGGTGGACAGT
     GGTATGAAAGAAAATGGGAAGGTGATATCATTTTTCAAACTAAAGGAGTCTCAACTGCCAGCTTTGGCAA
     TTTACCAGACTCTAGATGACGAGTGGGATACACTGCCCACAGCAGAAGTTTCCGTAGAGCATGTGCAAAA
     CTTTTGTGATGGATTCCTAAGTGGAAAATTGTTGAAAGAAAATCGTGAATCAGAAGGAAAGACTCCAAAG
60   GTGGAACTCTGACTTCTCCTTGGAACTACATATGGCCAAGTATCTACTTTATGCAAAGTAAAAAGGCACA
```

ACTCAAATCTCAGAGACACTAAACAACAGGATCACTAGGCCTGCCAACCACACACACACGCACGTGCACA
CACGCACGCACGCGTGCACACACACACGCGCACACACACACACACACACAGAGCTTCATTTCCTGTCT
TAAAATCTCGTTTTCTCTTCTTCCTTCTTTTAAATTTCATATCCTCACTCCCTATCCAATTTCCTTCTTA
TCGTGCATTCATACTCTGTAAGCCCATCTGTAACACACCTAGATCAAGGCTTTAAGAGACTCACTGTGAT

5   GCCTCTATGAAAGAGAGGCATTCCTAGAGAAAGATTGTTCCAATTTGTCATTTAATATCAAGTTTGTATA
CTGCACATGACTTACACACAACATAGTTCCTGCTCTTTTAAGGTTACCTAAGGGTTGAAACTCTACCTTC
TTTCATAAGCACATGTCCGTCTCTGACTCAGGATCAAAAACCAAAGGATGGTTTTAAACACCTTTGTGAA
ATTGTCTTTTTGCCAGAAGTTAAAGGCTGTCTCCAAGTCCCTGAACTCAGCAGAAATAGACCATGTGAAA
ACTCCATGCTTGGTTAGCATCTCCAACTCCCTATGTAAATCAACAACCTGCATAATAAATAAAAGGCAAT

10   CATGTTATAGGAGAAGGCTGGCATCTCTTTGCA

>DEX0367_016.nt.1
GCCTCCCCTGCTGACTGCTGTAGCTAGCCTTGCCCCATCCCACCCCTTCCCTCTCCATCCTCGCTCTTTT
CTCTCCTGGTTTATGGTTTGTCTCTTCCTCAAATGCAATTTGCAAAACTCTCCTGAACGAAACAACAGCT

15   TCTGGGAGGTGGTGGCAGCTGCAGGCACAGTGGCGCCGTGGGAAATCGTGAAAAACCCAGAATTCATTCT
TGGAGGGGCCACCAGGACTGATATCTGCCAGGGAGAGCTGGGAGACTGCTGGCTATTAGCCGCCATCGCC
TCCCTTACGCTTAATCAAAAAGCACTGGCCAGAGTCATCCCCCAGGACCAAAGCTTTGGCCCTGGTTATG
CCGGGATATTCCATTTCCAGTTCTGGCAGCACAGTGAGTGGCTGGACGTGGTGATCGATGACCGCCTGCC
CACCTTCAGGGACCGCTTGGTTTTCCTCCACTCTGCCGACCACAATGAGTTCTGGAGCGCCTTGCTGGAA

20   AAAGCCTACGCCAAGCTAAATGGGAGCTATGAAGCTCTGAAGGGAGGCAGCGCCATCGAGGCCATGGAAG
ACTTCACCGGGGGTGTGGCAGAGACCTTCCAAACTAAAGAGGCCCCCGAGAACTTCTATGAGATTCTAGA
GAAGGCTTTGAAGAGAGGCTCCCTGCTGGGCTGCTTCATTGATACCAGAAGTGCTGCAGAATCTGAGGCC
CGGACGCCGTTTGGTCTTATTAAGGGTCATGCCTACAGTGTAACGGGAATTGACCAGGTAAGCTTCCGAG
GCCAGAGAATCGAGCTCATCCGAATCCGGAACCCTTGGGGCCAGGTTGAGTGGAACGGGTCGTGGAGCGA

25   CAGGATGGCATTTAAGGACTTCAAGGCCCACTTTGATAAAGTGGAGATCTGCAACCTCACTCCCGATGCC
CTGGAGGAAGACGCGCATCCACAAATAAAATTGTCTCTGACTGAGAAAGATGAGGGGCAGGAGGAGTGTAG
TTTCCTTGTAGCCCTGATGCAGAAAGATAGAAGGAAACTCAAGAGATTTGGTGCCAATGTGCTGACAATC
GGCTATGCCATTTATGAGTGCCCTGACAAAGACGAACACCTGAACAAAGACTTCTTCAGATACCACGCTT
CTCGGGCCAGAAGCAAGACGTTCATCAACCTGAGAGAAGTCTCCGACCGGTTCAAGCTGCCCCCTGGGGA

30   GTACATCCTGATTCCCAGCACTTTTGAGCCCCACCAGGAAGCTGATTTCTGTCTGAGAATCTTTTCAGAG
AAAAAAGCCATTACCCGGGATATGGATGGAAATGTAGACATTGACCTTCCTGAGCCTCCAAAGCCAACTC
CACCTGACCAGGAGACAGAGGAGGAGCAGCGGTTTCGGGCTCTGTGCTGCAAAAGAAAAAGGACATCAAATTC
AAGAAGCTAAGCCTGATCTCCTGTAAAAAACATCATTTCCCTGATGGACACCAGCGGCAATGGGAAGCTGG

35   AGTTTGATGAATTCAAAGTGTTCTGGGACAAGCTGAAGCAGTGGATTAACCTTTTCCTTCGGTTTGATGC
TGACAAGTCCGGCACCATGTCTACCTATGAACTACGGACTGCACTGAAAGCTGCAGGCTTTCAGCTGAGC
AGCCACCTCCTGCAGCTGATTGTGCTCAGGTATGCGGATGAGGAGCTCCAGCTGGACTTCGATGACTTCC
TCAACTGCCTGGTCCGGCTGGAGAATGCGAGCCGGGTGTTCCAGGCTCTCAGTACAAAGAACAAGGAGTT
CATTCATCTCAATATAAATGAGTTCATCCATTTGACAATGAACATCTGAGGCTGCCTTGTAGAGATGCAG

40   CCTGCCCAGCTGAATCTTGGCTTCTGGACCTTGACCTTCAGAACTTCTCTTGGTGTGGAACCATTACGCC
CAGGGTTCACTCCCCTCTCATCGTCCGGCCTTCTCCCTTCATCTTGATCTGGGAAGAATGAAATGAACTC
AGCTACACTCTCTGATTTTGTGCTACTCCTTTGTAAAGTCACTGCCTTAAGGGGGCTGATGGCGCCACCT
GTGCCTTACATCCAGGTTCAGGCATCACTAGCTTTCCCACACTCTACTTTCCTTATTTCCTTCCATTAAG
AATTACTCAGAGTTCTAACGCACAGAATCCTGACTTCCATGTAGCTCCAGTCATTGTGATCAGACATCCT

45   TTATAAAACATGTTTTTATAAATGTGTATGTGGAATAAGATGAGGAGGAAAAGGAGTGAGGTCATAATTT
AGCATTTCTGCATCATAAGTGTGCTTCCTTATCAGGGAAGACATCAGGTAGTTCCGATAATTGGACAAGA
GT

>DEX0367_016.nt.2

50   TAATGAATGAGTGAATGGATGATGCCCTCACTCTTGCCCTCCTGTGCATCGTGTCTCTCCGCTTTGCTGC
AGGATGGCATTTAAGGACTTCAAGGCCCACTTTGATAAAGTGGAGATCTGCAACCTCACTCCCGATGCCC
TGGAGGAAGACGCGATCCACAAATGGGAGGTGACGGTCCATCAGGGAAGCTGGGTTCGCGGCTCCACGGC
TGGGGGCTGCCGCAATTTCCTGGATACCTTTTGGACCAATCCACAAATAAAATTGTCTCTGACTGAGAAA
GATGAGGGGCAGGAGGAGTGTAGTTTCCTTGTAGCCCTGATGCAGAAAGATAGAAGGAAACTCAAGAGAT

55   TTGGTGCCAATGTGCTGACAATCGGCTATGCCATTTATGAGTGCCCTGACAAAGACGAACACCTGAACAA
AGACTTCTTCAGATACCACGCTTCTCGGGCCAGAAGCAAGACGTTCATCAACCTGAGAGAAGTCTCCGAC
CGGTTCAAGCTGCCCCCTGGGGAGTACATCCTGATTCCTGATTCCAGCACTTTTGAGCCCCACCAGGAAGCTGATT
TCTGTCTGAGAATCTTTTCAGAGAAAAAAAGCCATTACCCGGGATATGGATGGAAATGTAGACATTGACCT
TCCTGAGCCTCCAAAGCCAACTCCACCTGACCAGGAGACAGAGGAGGAGCAGCGGTTTCGGGCTCTGTTT

60   GAACAAGTCGCTGGTGAGGACATGGAGGTGACAGCAGAGGAACTTGAGTATGTTTTAAATGCTGTGCTGC

AAAAGAAAAAGGACATCAAATTCAAGAAGCTAAGCCTGATCTCCTGTAAAAACATCATTTCCCTGATGGA
CACCAGCGGCAATGGGAAGCTGGAGTTTGATGAATTCAAAGTGTTCTGGGACAAGCTGAAGCAGTGGATT
AACCTTTTCCTTCGGTTTGATGCTGACAAGTCCGGCACCATGTCTACCTATGAACTACGGACTGCACTGA
AAGCTGCAGGCTTTCAGCTGAGCAGCCACCTCCTGCAGCTGATTGTGCTCAGGTATGCGGATGAGGAGCT
5 CCAGCTGGACTTCGATGACTTCCTCAACTGCCTGGTCCGGCTGGAGAATGCGAGCCGGGTGTTCCAGGCT
CTCAGTACAAAGAACAAGGAGTTCATTCATCTCAATATAAATGAGTTCATCCATTTGACAATGAACATCT
GAGGCTGCCTTGTAGAGATGCAGCCTGCCCAGCTGAATCTTGGCTTCTGGACCTTGACCTTCAGAACTTC
TCTTGGTGTGGAACCATTACGCCCAGGGTTCACTCCCCTCTCATCGTCCGGCCTTCTCCCTTCATCTTGA
TCTGGGAAGAATGAAATGAACTCAGCTACACTCTCTGATTTTGTGCTACTCCTTTGTAAAGTCACTGCCT
10 TAAGGGGGCTGATGGCGCCACCTGTGCCTTACATCCAGGTTCAGGCATCACTAGCTTTCCCACACTCTAC
TTTCCTTATTTCCTTCCATTAAGAATTACTCAGAGTTCTAACGCACAGAATCCTGACTTCATGTAGCTC
CAGTCATTGTGATCAGACATCCTTTATAAAACATGTTTTTATAAATGTGTATGTGGAATAAGATGAGGAG
GAAAAGGAGTGAGGTCATAATTTAGCATTTCTGCATCATAAGTGTGCTTCCTTATCAGGGAAGACATCAG
GTAGTTCCGATAATTGGACAAGAGT

15 >DEX0367_017.nt.1
GCCCTTAGCGTGGTCGCGGCCGAGGTACGTGTAGTGGGCCGGAATGGAGAATCCAGTGAACTGGACCTAC
AAGGCATCCGAATCGACTCAGATATTAGCGGCACCCTCAAGTTTGCGTGTGAGAGCATTGTGGAGGAATA
CGAGGATGAACTCATTGAATTCTTTTCCCGAGAGGCTGACAATGTTAAAGACAAACTTTGCAGTAAGCGA
20 ACAGATCTTTGTGACCATGCCCTGCACATATCGCATGATGAGCTATGAACCACTGGAGCAGCCCACACTG
GCTTGATGGATCACCCCCAGGAGGGGAAAATGGTGGCAATGCCTTTTATATATTATGTTTTTACTGAAAT
TAACTGAAAAAATATGAAACCAAAAGT

>DEX0367_018.nt.1
25 GGTAATAGAGCTTCGACTTCGACGGCCGCCATTTTCTTTCTTTCTTAGCTGTTAGCTGAGAGGAAGTCTC
TGAACAGGCGGCAGCGGCTCTTATAGTGCAACCATGGCAGACTATTCAACAGTGCCTCCCCCCTCTTCTG
GCTCAGCTGGTGGCGGTGGTGGCGGCGGTGGTGGTGGAGGAGTTAACGACGCTTTCAAAGATGCACTGCA
GAGAGCCCGGCAGATTGCAGCAAAAATTGGAGGTGATGCAGGGACATCACTGAATTCAAATGACTATGGT
TATGGGGGACAAAAAAGACCTTTAGAAGATGGAGATCAACCAGATGCTAAGAACTGTAATGACAGAAGAATACAA
30 ACTCTTTTGGAACACAGTTACCACCGATGCATCAGCAGCAAAGCAGAGGTGAACAGATCTCACGCATACAACAGGAA
AGTTCCAGATGGAATGGTTGGATTCATAATTGGCAGAGGACTGGTGGCCTTCCAGAAAGGTCCTGTATGTTAACTGGAA
TCTGGATGCAAAATACAGATAGCTCCTGACAGTGGTGGCCTTCCAGAAAGGTCCTGTATGTTAACTGGAA
CACCTGAATCTGTCCAGTCAGCAAAACGGTTACTGGACCAGATTGTTGAAAAAGGAAGACCAGCTCCTGG
CTTCCATCATGGCAGGAGACCGGGAAATGCAGTTCAAGAAATCATGATTCCAGCTAGCAAGGCAGGATTA
GTCATTGGAAAAGGGGGAGAAACTATTAAACAGCTTCAGGAACGGGCTGGAGTTAAAATGGTTATGATTC
35 AAGACGGGGTAGCGGCCGCCGCGGCCGTATTTTTCACATCCATCATAAGGTTGGCATTCAATAAGGAAAA
AGTTCTAACTCCAGTATTTAAATTGTACATAAATCCCAAATGTTCTTAAAGAACACTCAGGGACATGTTTG
TTGCCTGGCATTGGTAATGAAAGGTTGGTTTTTGAAACTTGAAATTTCACCATTGGTTTTTTTCCTATCA
TTTCTGCATATCCAGCAAAAGGAATCTCATGTTGACTCCTGGCAGAGTTCAGTGGCTTCAGTCTGTCTAT
40 CTGCGG

>DEX0367_019.nt.1
ACAAAGTTAAATTTCAACACATTCTCTATGCTAGTGTGACAAAAGCAGCCCCATAATTTGGTTTTTATTG
TTGACCTTTACAGGATGAAGGAGACACCACAATGGCTGAGCACTTCCTGACGTTGCTGGTAGTGCCTGCC
ATCAAGAAAGATTATGGTTCCCAGGAAGACTTCACTCAAGTGTGGAACACCACCATGAAAGGGCTCAAGT
45 GCTGTGGCTTCACCAACTATACGGATTTTGAGGACTCACCCTACTTCAAAGAGAACAGTGCCTTTCCCCC
ATTCTGTTGCAATGACAACGTCACCAACACAGCCAATGAAACCTGCACCAAGCAAAAGGCTCACGACCAA
AAAGTAGAGGGTTGCTTCAATCAGCTTTTGTATGACATCCGAACTAATGCAGTCACCGTGGGTGGTGTGG
CAGCTGGAATTGGGGGCCTCGAGCTGGCTGCCATGATTGTGTCCATGTATCTGTACCTGCCCGGGCGGCC
50 AGCGTGGTCGCGGCCGAGGTACAGCGACAGCGTCGGCAGGGTCAGCTTCCCCAGCCTGGTCTGCTTCCTG
ATGCGGCTTGAAGCCATGGCAAAGACCTTCCGCAACTCTCTAAGGATGGAAAAGGACTCTACCTGACAGA
AATGGAGTGGATGAGCCTGGTCATGTCTGGTTGCCAGGAGGAGCAAGCAAAGTCAGCAGGACATGAAGTT
GTATTAAATGGACTTCGTGATTTTGTTTTGCACTAAAGTTTCTGTGATTTAACAATTAAATTCTGTTAGC
AC

55 >DEX0367_019.nt.2
GTGGAGGGCACGCAGACATTCTGGGAAGCCACTTGCCCCACCCCTGGGCTGCTTCTTCTTGAGATCAGGA
GGGGCGTTGCCCAGGGCTGGTGTTGCCAGGTGGAGGCCTGCTGAGGCAGTGGTTGTGGGGATCGGTCTCC
AGGCAGCAGGGGGCAGCAGGGTCAAGGAGAGGCTAACTGGCCACGGGTGGGGCCAGCAGGCGGGCAGAAG
60 GAGGCTTTAAAGCGCCTACCCTGCCTGCAGGTGAGCAGTGGTGTGTGAGAGCCAGGCGTCCCTCTGCCTG

CCCACTCAGTGGCAACACCCGGGAGCTGTTTTGTCCTTTGTGGAGCCTCAGCAGTTCCCTCTTTCAGAAC
TCACTGCCAAGAGCCCTGAACAGGAGCCACCATGCAGTGCTTCAGCTTCATTAAGACCATGATGATCCTC
TTCAATTTGCTCATCTTTCTGTGTGGTGCAGCCCTGTTGGCAGTGGGCATCTGGGTGTCAATCGATGGGG
CATCCTTTCTGAAGATCTTCGGGCCACTGTCGTCCAGTGCCATGCAGTTTGTCAACGTGGGCTACTTCCT
CATCGCAGCCGGCGTTGTGGTCTTTGCTCTTGGTTTCCTGGGCTGCTATGGTGCTAAGACTGAGAGCAAG
TGTGCCCTCGTGACGTTCTTCTTCATCCTCCTCCTCATCTTCATTGCTGAGGTTGCAGCTGCTGTGGTCG
CCTTGGTGTACACCACAATGGCTGAGCACTTCCTGACGTTGCTGGTAGTGCCTGCCATCAAGAAAGATTA
TGGTTCCCAGGAAGACTTCACTCAAGTGTGGAACACCACCATGAAAGGGCTCAAGTGCTGTGGCTTCACC
AACTATACGGATTTTGAGGACTCACCCTACTTCAAAGAGAACAGTGCCTTTCCCCCATTCTGTTGCAATG
ACAACGTCACCAACACAGCCAATGAAACCTGCACCAAGCAAAAGGCTCACGACCAAAAAGTAGAGGGTTG
CTTCAATCAGCTTTTGTATGACATCCGAACTAATGCAGTCACCGTGGGTGGTGTGGCAGCTGGAATTGGG
GGCCTCGAGTTCTTTTCCAACTCAGCTCGAAGGCCACCTCTTCCAGAAAGCCTCTATAGCACTCCCATCA
GAAGAGATCACGTCTTTCTACAACCCTCCCCTCCATGACTTTCATGGCTCTTAGAGCCTCTGCTGTGTCT
GCTTCATCCTGGAAGTATCACAATCCTCCACCACACTGAACCCCTCAAGGTAGGGCCAGGTCTGATTACT
TTCAGGTCCCCAGTGCCCAGCACAAGGCTGAGGCCAAAGAAAGGACCAGGGGATGGTTATAAAATAAATC
AATGAATTGACT

>DEX0367_020.nt.1
CGGTGGCCCCTCCCGCCCACGCGCCAGAGTCGCAGTGGGCGGGCCTACGTGCTCCGCCCGCTGTGAGCCT
GTCCGGCCCCGCCCGCTCCGGAGCAACCCGCGAGCTTACACCGGCTTCTCTCTGTCCTCAGCCCGCGCG
CCGCCATCGCCCGTCATGCTGGGCGCCGCCTCTCCGCCGCTGCGCTGTGGCCGCAACCACCCGGGCCGACCC
TCGAGGCCTCCTGCTCACTCCGCCCGGACCCCCGGCCCCGCCGTGGCTATCCAGTCAGTTCGCTGCTATTCC
CATGGGTCACAGGAGACAGATGAGGAGTTTGATGCTCGCTGGGTAACATACTTCAACAAGCCAGATATAG
ATGCCTGGGAATTGCGTAAAGGGATAAACACACTTGTTACCTATGATATGGTTCCAGAGCCCACAGTTCG
TATCCTAGAGGTTGTTAAGGACAAAGCAGGACCTCATAAGGAAATCTACCCCTATGTCATCCAGGAACTT
AGACCAACTTTTAAATGAACTGGGAATCTCCACTCCGGAGGAACTGGGCCTTGACAAAGTGTAAACCGCAT
GGATGGGCTTCCCCAAGGATTTATTGACATTGCTACTTGAGTGTGAACAGTTACCTGGAAATACTGATGA
TAACATATTACCTTATTTGAACAAGTTTTCCTTTATTGAGTACCAAGCCATGTAATGGTAACTTGGACTT
TAATAAAAGGGAAATGAGTTTGAACTGGCMCMCACMCCCCCCACCCCCCAACACACAACACGGCGGACCT
AATAGAAAGACACCATAAGGGGGCCCCAATTTACGTCTACCCGAATATTACTTACAATCTGGCCCCTTCT
GCGGAACATATACACCCAGCTAGTGCCTCCCTTCAACCCTTACACGGAAAACG

>DEX0367_021.nt.1
GGGCTTTTTCGAGCCCATAGGTCGGGACAGCCCAGGACTGGGCATGGGAGGCTTGGAGTGGCTCTGGGTA
GGGGCCCAGCCAGTGCCCTAGGGGACTTCATATACCCTCTCTTCCTTTGCAGGAAACCCCCTCGGCTTCA
GGACTTCCGCTGCTTAGCTGTGCTGGGCCGGGGACACTTTGGGAAGGTAGTGGGCTGAAGAGGGTGGTAT
GGGACAGGATTGGGGGCCTCATCACATGAGCCGGGAGGACCTGATCTGTGGGGGCGGTAAAGGCTCCCTG
GTCTGGCCCACCTGGAGCACAACCTCTCTCTGGCCCCAGGTCCTCCTGGTCCAGTTCAAGGGGACAGGGA
AATACTACGCCATCAAAGCACTGAAGAAGCAGGAGGTGCTCAGCCGGGACGAGATAGAGAGCCTGTACTG
CGAGAAGCGGATCCTGGAGGCTGTGGGCTGCACAGGGCACCCTTTCCTGCTCTCCCTCCTTGCCTGCTTC
CAGACCTCCAGCCATGCCTGCTTTGTGACTGAGTTTGTGCCTGGTGGTGACCTCATGATGCAGATCCACG
AGGATGTCTTCCCCGAGCCCCAGGCCCGCTTCTACGTGGCTTGTGTTGTCCTGGGGCTGCAGTTCTTACA
CGAGAAGAAGATCATTTACAGGGACCTGAAGTTGGATAACCTTCTGCTGGATGCCCAGGGATTCCTGAAG
ATCGCAGACTTTGGACTCTGCAAGGAAGGGATCGGCTTCGGGGACCGGACTAGCACCTTCTGTGGCACCC
CGGAGTTCCTGGCTCCCGAGGTGCTGACCCAGGAGGCATACACACGGGCTGTGGACTGGTGGGGGCTGGG
TGTGCTGCTCTACGAGATGCTGGTGGGTGAGTGCCCGTTCCCAGGGGACACAGAGGAAGAGGGTGTTTGAC
TGCATCGTCAACATGGACGCCCCCTACCCCGGCTTTCTGTCGGTGCAAGGGCTTGAGTTCATTCAGAAGC
TCCTCCAGAAGTGCCCGGAGAAGCGCCTCGGGGCAGGTGAGCAGGATGCCGAGGAGATCAAGGTCCAGCC
ATTCTTCAGGACCACCAACTGGCAAGCCCTGCTCGCCCGCACCATCCAGCCCCCCTTCGTGCCTACCCTG
TGTGGCCCTGCGGACCTGCGCTACTTTGAGGGCGAGTTCACAGGGCTGCCGCCTGCCCTGACCCCACCTG
CACCCCACAGCCTCCTCACTGCCCGCCAACAGGCCGCCTTCCGGGACTTCGACTTTGTGTCAGAGCGATT
CCTGGAACCCTGAGGGCATCTCCTGGCACCTCTGTCCCCTTCCCCCACAGACTGTTAGAGCCTCTGCTCG
TTCACCCGTGCGCCCTGCCTGGAGGTCCAGGCCTTGCTGGGTACTTCTGAGCCCTTGGGATTCAAAGTGG
CAGCCATGGGGCCACTGTTGTGGGCTTTGCTCAGTGTCACTGGGCAAAGTGTGTCCCTTCCCCCTCCAGC
TCGCCCTCTTCTACCTCCCAGCGAGACCTGGCCCAGAAAGGGTGCCGCAGCAAGGAGTGATATGGTTTGT
CTTTTTAAGACTGGACTTGCTTTATATTAAATTTGTAAAAGTGAAAAAAAAAG

>DEX0367_021.nt.2
ACCGCACCTGGCCTAGTATACATTCTTGTAGTGACTAATATCTAGTCTGGGCCAACCACTCCGAGGAGAG
AGACAATTTTGGATTAGCCCACATTACACCAACAACGTGCGAGACAAAGAGAGACTTGAGGGGCTTGGTC

```
GAGGAGGCCCAGCGCGTTAGAGGCGGGAGCCAGGTTGTAGAGGGGCTCAAAGTTAGAGGAGGCGGCCAGC
TCCACCGGCTGCACCCGGCTGGGCGACCCGGCAACCTTCAGGGGGCGCTGTTTCCCATCCCTGTCCCCT
AACTGGCCGGCTCCCGCGGGCGCGCGGCGGGGAAGGCCAGAGGACCTGGGCGCGGGCGATGTGCCTCCTG
AGCGTCCAAACCGGGGGTGAGGCGCGGTCACGCCCAGCGGGAACCGCAGGCGCCGAAGCCCGGGTACTGG
GCCCAGAATCCCGCGGAATTTTGGATCCGAGGGAGGCGCTGGGGCGCGGGGACCTCGGGCGTGGGGTCCCG
GGCGCTGGATCGGCGCGGACGGGAGGCGGCGCTGGTCCCGCGGGCCAGCGGGTCTCGGGAGGGGGCGCCC
GATCCCGCGTCTCCGGCGCCGCTTCCCGGGAAGTTTCAAGTTTGAAAGTCCTGGCGGAGGGTCTGCGGCT
TCCGGGACCGGAGTGGCTGAGAGGAGGGCCCCAAGCGGCCGGAGCGGCGCCATGGAGGAGGGGGCGCCGC
GGCAGCCTGGGCCGAGCCAGTGGCCCCCAGAGGATGAGAAGGAGGTGATCCGCCGGGCCATCCAGAAAGA
GCTGAAGATCAAGGAGGGGGTGGAGAACCTGCGGCGCGTGGCCACAGACCGCCGCCACTTGGGCCATGTG
CAGCAGCTGCTGCGGTCCTCCAACCGCCGCCTGGAGCAGCTGCATGGCGAGCTGCGGGAGCTGCACGCCC
GAATCCTGCTGCCCGGCCCTGGGCCTGGCCCAGCTGAGCCTGTGGCCTCAGGACCCCGGCCGTGGGCAGA
GCAGCTCAGGGCTCGGCACCTAGAGGCTCTCCGGAGGCAGCTGCATGTGGAGCTGAAGGTGAAGCAGGGG
GCTGAGAACATGACCCACACGTGCGCCAGTGGCACCCCCAAGGAGAGGAAGCTCCTGGCAGCTGCCCAGC
AGATGCTGCGGGACAGCCAGCTGAAGGTGGCCCTGCTGCGGATGAAGATCAGCAGCCTGGAGGCCAGTGG
GTCCCCGGAGCCAGGGCCTGAGCTGCTGGCGGAGGAGCTACAGCATCGACGTTGAGGCAGCTGTG
GCTGAGGGCGCCAAGAACGTGGTGAAACTGCTTAGTAGCCGGAGAACACAGGACCGCAAGGCACTGGCTG
AGGCCCAGGCCCAGCTACAGGAGTCCTCTCAGAAACTGGACCTCCTGCGCCTGGCCTTGGAGCAGCTGCT
GGAGCAACTGCCTCCTGCCCACCCTTTGCGCAGCAGAGTGACCCGAGAGTTGCGGGCTGCGGTGCCTGGA
TACCCCCAGCCTTCAGGGACACCTGTGAAGCCCACCGCCCTAACAGGGACACTGCAGGTCCGCCTCCTGG
GCTGTGAACAGTTGCTGACAGCCGTGCCTGGGCGCTCCCCAGCGGCCGCACTGGCCAGCAGCCCCTCCGA
GGGCTGGCTTCGGACCAAGGCCAAGCACCAGCGTGGCCGAGGCGAGCTTGCCAGCGAGGTGCTGGCTGTG
CTAAAGGTGGACAACCGTGTTGTGGGGCAGACGGGCTGGGGGCAGGTGGCCGAACAGTCCTGGGACCAGA
CCTTTGTCATCCCACTGGAGCGAGCCCGTGAGCTGGAGATTGGGGTACACTGGCGGGACTGGCGGCAGCT
ATGTGGCGTGGCCTTCCTGAGACTTGAGGACTTCCTGGACAATGCCTGTCACCAACTGTCCCTCAGCCTG
GTACCGCAGGGACTGCTTTTTGCCCAGGTGACCTTCTGCGATCCTGTCATTGAGAGGCGGCCCCGGCTGC
AGAGGCAGGAACGCATCTTCTCTAAACGCAGAGGCCGCCTTCCGGGGACTTCGACTTTGTGTCAGAGCGAT
TCCTGGAACCCTGAGGGCATCTCCTGGCACCTCTGTCCCCTTCCCCCACAGACTGTTAGAGCCTCTGCTC
GTTCACCCGTGCGCCCTGCCTGGAGGTCCAGGCCTTGCTGGGTACTTCTGAGCCCTTGGGATTCAAAGTG
GCAGCCATGGGGCCACTGTTGTGGGCTTTGCTCAGTGTCACTGGGCAAAGTGTGTCCCTTCCCCCTCCAG
CTCGCCCTCTTCTACCTCCCAGCGAGACCTGGCCCAGAAAGGGTGCCGCAGCAAGGAGTGATATGGTTTG
TCTTTTTTAAGACTGGACTTGCTTTATATTAAATTTGTAAAAGTGAAAAAAAAAAG
```

```
>DEX0367_022.nt.1
GAGCAAGCTGGTGTCAGAGCCTGGACCTACAGCGCTGTTGGTGGAGGTCCTGCCTCCAGGAATAGATGGA
CATGGCCTGGCAGATGATGCAGCTGCTGCTTCTGGCTTTGGTGACTGCTGCGGGGAGTGCCCAGCCCAGG
AGTGCGCGGGCCAGGACGGACCTGCTCAATGTCTGCATGAACGCCAAGCACCACAAGACACAGCCCAGCC
CCGAGGACGAGCTGTATGGCCAGGTGGGAGCTCCTCAAGGGCCCTCCCCAGGAAGTGTTCCTCTGGATGA
CCTACCTGGGGCAGAGGAGCCAGAATATGGAGGAGATGGCCTGTGTGGTGGGAGAGACTTAGTCCTGTGTCT
TCCCCACCCAGTGCAGCCCTGGAAGAAGAATGCCTGCTGCACGGCCAGCACCAGCCAGGAGCTGCACAA
GGACACCTCCCGCCTGTACAACTTTAACTGGGATCACTGTGGTAAGATGGAACCCACCTGCAAGCGCCAC
TTTATCCAGGACAGCTGTCTCTGAGTGCTCACCCAACCTGGGGCCCTGGATCCGGCAGGTCAACCAGAGC
TGGCGCAAAGAGCGCATTCTGAACGTGCCCCTGTGCAAAGAGGACTGTGAGCGCTGGTGGGAGGACTGTC
GCACCTCCTACACCTGCAAAAGCAACTGGCACAAAGGCTGGAATTGGACCTCAGGGATTAATGAGTGTCC
GGCCGGGGCCCTCTGCAGCACCTTTGAGTCCTACTTCCCCACTCCAGCCGCCCTTTGTGAAGGCCTCTGG
AGCCACTCCTTCAAGGTCAGCAACTATAGTCGAGGGAGCGGCCGCTGCATCCAGATGTGGTTTGACTCAG
CCCAGGGCAACCCCAATGAGGAGGTGGCCAAGTTCTATGCTGCGGCCATGAATGCTGGGGCCCCGTCTCG
TGGGATTATTGATTCCTGATCCAAGAAGGGTCCTCTGGGGTTCTTCCAACAACCTATTCTAATAGACAAA
TCCACATG
```

```
>DEX0367_023.nt.1
CCAAGCCCACGCGGTTCCGGGTCCCAAGATTTGGCTCCAAGAGCTATTCGCAGAGAAACAGCCTCGCACA
AAAAGCCACCGCCCCCGAAACCCGCGGGAAATCTGTTTCCTGTGAAAGGCCTGGCCTCCGGGCATTGGCG
GTTTCCCTGATCCTCTCTCCCAAATTTTCCGGCAGGGGATTGATCGAGGCGACATTTAGGACTGGGGTGG
GCGAGGGAGTGGCGGTTCTCGTCCCAAGTTACAGGTGAGGACGTTCAGGCCAAGGCGTAGCCAAGATCGC
GCCGCCCTGTAGCTTGAGGAGGCGCCGGGATTCGAACCAAGGCTGTGCCCAATCCTAAGCCCTGCGAGGT
CTTGTGGAACCGACATGACCCACTCTTTGGTTTGTCCAGAGACAGTGAGCAGGGTGAGTTCAGTGCTGAA
TCGCAACACTCGGCAGTTTGGAAAAAAACATCTTTTCGACCAGGATGAGGAGACATGTTGGAACTCAGAC
CAGGGCCCCTCCCAGTGGGTGACGCTGGAGTTTCCCCAGCTCATCCGTGTCTCCCAGCTGCAGATCCAGT
TTCAGGGTGGCTTCTCCAGTCGCCGGGGCTGCCTGGAAGGTTCACAGGGCACTCAGGCTCTCCACAAGAT
```

```
TGTAGATTTCTACCCTGAGGACAACAACTCGCTTCAGGATATCCTGCCCCTGCTTCTGGGGTGGGTGGTC
TGTGCCTCTGGAAGCTTTGGACTCACCCAAACCCGAATGGGAACCCTGATGTGTGATCTTGAACACGTGA
CCTCTAGGACTCCACTTTTTTCTGTAAAATGGTGCTGATGACCACCTGTTCTGTGTGAGAGGCTTAGAGGG
ACCTGCAGCTTCCCCAGGTTGCCCACATCGGTGGAGTGTGAGCTGGGTCTGGAAGGGGTGGGGGCCTGCG
GGTGGGAAACCAGAAATCACCAGTTTCCCTTGACTGCCCACGCTCACACACTTTCCCCATACCAGCTGCT
GAAGTGGACCGGCTGAAGGTGACGTTTGAGGATGCCACTGACTTTTTTGGCCGTGTGGTCATCTACCACC
TGCGGGTGCTTGGGGAGAAGGTGTGAGACCTCTAGGGGCTGTCTCCTCCAGGAAGCCCTCCGGGAAGCAC
AGCAAAGTCCCTCATTCTGCACAGAAGGTTTATTGGTTCCTCTTGGGAAGGGTCCCCTCCCACCACCTGT
CCAGAAGCTGCCTTTGAAGTCAGTTCTGGGTTTCCCCAGCTCTGGCTGACCATTTTGTTCCCTGAGTGTC
TGAGTCCCGGCAGGCGGCTTCACTCAGGGTCAGCGGGCACCAGGTTGCTCTGGAAGAGCTTGAGGATGTG
GTTCTCGATCACCTGTTGCACTGAGATGGGGCAGGGAAAAGCGTGGGCTGTGAGCTTGAATCGGGAGTGG
GGTGGAGGCACAGGCCAACCTGCGCTCTCCCCTTAGGGGACAAACAGGGACCCTTGCAGAGACCTGCATT
ACAGAGCAAAGCTGGGAGAACCGAGGACTCACCCCAGGAACCTCAACTTCCCCTAAGTTTGTAAAAATGC
ACTTTTGTAGCTTTTCAGAATTAATATAAAAATATTTCCTACCAAGCAAAA
```

>DEX0367_023.nt.2
```
CCAAGCCCACGCGGTTCCGGGTCCCAAGATTTGGCTCCAAGAGCTATTCGCAGAGAAACAGCCTCGCACA
AAAAGCCACCGCCCCGAAACCCGCGGGAAATCTGTTTCCTGTGAAAGGCCTGGCCTCCGGGCATTGGCG
GTTTCCCTGATCCTCTCTCCCAAATTTTCCGGCAGGGGATTGATCGAGGCGACATTTAGGACTGGGGTGG
GCGAGGGAGTGGCGGTTCTCGTCCCAAGTTACAGGTGAGGACGTTCAGGCCAAGGCGTAGCCAAGATCGC
GCCGCCCTGTAGCTTGAGGAGGCGCCGGGATTCGAACCAAGGCTGTGCCCAATCCTAAGCCCTGCGAGGT
CTTGTGGAACCGACATGACCCCACTCTTTGGTTTGTCCAGAGACAGTGAGCAGGGTGAGTTCAGTGCTGAA
TCGCAACACTCGGCAGTTTGGAAAAAAAACATCTTTTCGACCAGGATGAGGAGACATGTTGGAACTCAGAC
CAGGTTCACAGGGCACTCAGGCTCTCCACAAGATTGTAGATTTCTACCCTGAGGACAACAACTCGCTTCA
GACTTTCCCCATACCAGCTGCTGAAGTGGACCGGCTGAAGGTGACGTTTGAGGATGCCACTGACTTTTTT
GGCCGTGTGGTCATCTACCACCTGCGGGTGCTTGGGGAGAAGGTGTGAGACCTCTAGGGGCTGTCTCCTC
CAGGAAGCCCTCCGGGAAGCACAGCAAAGTCCCTCATTCTGCACAGAAGGTTTATTGGTTCCTCTTGGGA
AGGGTCCCCTCCCACCACCTGTCCAGAAGCTGCCTTTGAAGTCAGTTCTGGGTTTCCCCAGCTCTGGCTG
ACCATTTTGTTCCCTGAGTGTCTGAGTCCCGGCAGGCGGCTTCACTCAGGGTCAGCGGGCACCAGGTTGC
TCTGGAAGAGCTTGAGGATGTGGTTCTCGATCACCTGTTGCACTGAGATGGGGCAGGGAAAAGCGTGGGC
TGTGAGCTTGAATCGGGAGTGGGGTGGAGGCACAGGCCAACCTGCGCTCTCCCCTTAGGGGACAAACAGG
GACCCTTGCAGAGACCTGCATTACAGAGCAAAGCTGGGAGAACCGAGGACTCACCCCAGGAACCTCAACT
TCCCCTAAGTTTGTAAAAATGCACTTTTGTAGCTTTTCAGAATTAATATAAAAATATTTCCTACCAAGCA
AAA
```

>DEX0367_024.nt.1
```
ACAGGTGCCCATGCCTCTGGCTGCATTGGTGAGGATCCCCAGGGCATACCCTGCCCCTTGATGAGGCCCA
CCCCACGGGTGGTTGTACCAACCCTTACGGCAAGTCCAAGTTCTTCATCGAGGAAATGATCCGGGACCTG
TGCCAGGCAGACAAGGGGTGGACAGCAGCCTTAGGGCTGGACAGGATGTGTGAGGATCTCTGGCGCTGGC
AGAAGCAGAATCCTTCAGGCTTTGGCACGCAAGCCTGAGGACCCTCCCCTACCAAGGACCAGGAAAAGCA
GCAGCTGCCTGCTCTCCAGCCTCTGGCAGGAACTCAGGGCCCTGGAGCTGCTGGGGCCAAGCCAAGGCCT
CCCCTACCTCAAACCCCAGCTGGGCCCGCTTAGCCCACCAGGCATGAGGCCAAGGGCTCCACTGACCAGG
AGGCCGAGGTCTCTAACTCTTATCTTCCACAGGGTCCAAGAGTTCATCAGGACCCCCAAGAGTGAGTGAG
GGGGCAAGGCTCTGGCACAAAAACCTCCTCCTCCCAGGCACTCATTTATATTGCTCTGAAAGAGCTTTCCA
AAGTATTTAAAAATAAAAACAAGTTTTCTTACACTGGG
```

>DEX0367_024.nt.2
```
CGTGGAGGGGCGGGGTCGGGGCCGGAGTGGGCCGTCAGCACTTAAAGGGCCCGCGGCTCGGGCGTAGGAG
GCGGTGCCTCTGCAGCAAGCGTGGGCGCGGGAACCCGAGCAGGACTCTCCAGTCCTCAGTCACCTTGGA
CAAAGAAGTGTGGATCCTCAGATTCCATCTTTTCCAACTCCAAGGTGCCATGGCAGAGAAGGTGCTGGTA
ACAGGTGGGGCTGGCTACATTGGCAGCCACACGGTGCTGGAGCTGCTGGAGGCTGGCTACTTGCCTGTGG
TCATCGATAACTTCCATAATGCCTTCCGTGGAGGGGGCTCCCTGCCTGAGAGCCTGCGGCGGGTCCAGGA
GCTGACAGGCCGCTCTGTGGAGTTTGAGGAGATGGACATTTTGGACCAGGGAGCCCTACAGCGTCTCTTC
AAAAAGTACAGCTTTATGGCGGTCATCCACTTTGCGGGGCTCAAGGCCGTGGGCGAGTCGGTGCAGAAGC
CTCTGGATTATTACAGAGTTAACCTGACCGGGACCATCCAGCTTCTGGAGATCATGAAGGCCCACGGGGT
GAAGAACCTGGTGTTCAGCAGCTCAGCCACTGTGTACGGGAACCCCCAGTACCTGCCCCTTGATGAGGCC
CACCCCACGGGTGGTTGTACCAACCCTTACGGCAAGTCCAAGTTCTTCATCGAGGAAATGATCCGGGACC
TGTGCCAGGCAGACAAGACTTGGAACGCAGTGCTGCTGCGCTATTTCAACCCCACAGGTGCCCATGCCTC
TGGCTGCATTGGTGAGGATCCCCAGGGCATACCCAACAACCTCATGCCTTATGTCTCCCAGGTGGCGATC
GGGCGACGGGAGGCCCTGAATGTCTTTGGCAATGACTATGACACAGAGGATGGCACAGGTGTCCGGGATT
```

ACATCCATGTCGTGGATCTGGCCAAGGGCCACATTGCAGCCTTAAGGAAGCTGAAAGAACAGTGTGGCTG
CCGGGTAGGAAGAGAAGGGAGAAGTGAGGGAGGGGAGGGACCAGATCCAGGCAGAGCGGCCCAGAGGCGT
GGTCAGTCCAGCCCGCTCCATAAACCCTGTTCTCCGTGGGCCAGATCTACAACCTGGGCACGGGCACAGG
CTATTCAGTGCTGCAGATGGTCCAGGCTATGGAGAAGGCCTCTGGGAAGAAGATCCCGTACAAGGTGGTG
GCACGGCGGGAAGGTGATGTGGCAGCCTGTTACGCCAACCCCAGCCTGGCCCAAGAGGAGCTGGGGTGGA
CAGCAGCCTTAGGGCTGGACAGGATGTGTGAGGATCTCTGGCGCTGGCAGAAGCAGAATCCTTCAGGCTT
TGGCACGCAAGCCTGAGGACCCTCCCCTACCAAGGACCAGGAAAAGCAGCAGCTGCCTGCTCTCCAGCCT
CTGGCAGGAACTCAGGGCCCTGGAGCTGCTGGGGCCAAGCCAAGGCCTCCCCTACCTCAAACCCCAGCTG
GGCCCGCTTAGCCCACCAGGCATGAGGCCAAGGGCTCCACTGACCAGGAGGCCGAGGTCTCTAACTCTTA
TCTTCCACAGGGTCCAAGAGTTCATCAGGACCCCCAAGAGTGAGTGAGGGGGCAAGGCTCTGGCACAAAA
CCTCCTCCTCCCAGGCACTCATTTATATTGCTCTGAAAGAGCTTTCCAAAGTATTTAAAAAATAAAAACAA
GTTTTCTTACACTGGG

>DEX0367_024.nt.3
CGTGGAGGGGCGGGGTCGGGGCCGGAGTGGGCCGTCAGCACTTAAAGGGCCCGCGGCTCGGGCGTAGGAG
GCGGTGCCTCTGCAGCAAGCGTGGGGCGCGGGAACCCGAGCAGGACTCTCCAGTCCTCAGTCACCTTGGA
CAAAGAAGTGTGGATCCTCAGATTCCATCTTTTCCAACTCCAAGGTGCCATGGCAGAGAAGGTGCTGGTA
ACAGGTGGGGCTGGCTACATTGGCAGCCACACGGTGCTGGAGCTGCTGGAGGCTGGCTACTTGCCTGTGG
TCATCGATAACTTCCATAATGCCTTCCGTGGAGGGGGCTCCCTGCCTGAGAGCCTGCGGCGGGTCCAGGA
GCTGACAGGCCGCTCTGTGGAGTTTGAGGAGATGGACATTTTGGACCAGGGAGCCCTACAGCGTCTCTTC
AAAAAGTACAGCTTTATGGCGGTCATCCACTTTGCGGGGCTCAAGGCCGTGGGCGAGTCGGTGCAGAAGC
CTCTGGATTATTACAGAGTTAACCTGACCGGGACCATCCAGCTTCTGGAGATCATGAAGGCCCACGGGGT
GAAGAACCTGGTGTTCAGCAGCTCAGCCACTGTGTACGGGAACCCCCAGTACCTGCCCCTTGATGAGGCC
CACCCCACGGGTGGTTGTACCAACCCTTACGGCAAGTCCAAGTTCTTCATCGAGGAAATGATCCGGGACC
TGTGCCAGGCAGACAAGACTTGGAACGCAGTGCTGCTGCGCTATTTCAACCCCACAGGTGCCCATGCCTC
TGGCTGCATTGGTGAGGATCCCCAGGGCATACCCAACAACCTCATGCCTTATGTCTCCCAGGTGGCGATC
GGGCGACGGGAGGCCCTGAATGTCTTTGGCAATGACTATGACACAGAGGATGGCACAGGTGTCCGGGATT
ACATCCATGTCGTGGATCTGGCCAAGGGCCACATTGCAGCCTTAAGGAAGCTGAAAGAACAGTGTGGCTG
CCGGGTAGGAAGAGAAGGGAGAAGTGAGGGAGGGGAGGGACCAGATCCAGGCAGAGCGGCCCAGAGGCGT
GGTCAGTCCAGCCCGCTCCATAAACCCTGTTCTCCGTGGGCCAGATCTACAACCTGGGCACGGGCACAGG
CTATTCAGTGCTGCAGATGGTCCAGGCTATGGAGAAGGCCTCTGGGAAGAAGGTAGGCCCCCCCACCCCA
CCCACCTCACCCACCTCACCCCCACCCAGCTCTCAGCAACAGAGCAGCTTTGGCCCTTCCCACTCACCTCT
CTGGAGGGTACCTGGCACTGCCCAGTCTCCTCAGCCTGCCTAGCACCCGGTGCCTGCGGGCCAGCCGCTG
CTCTGCACTTCTGCTTCTTAAAGACCTGGCCAGCAGCTGGGCCAGGGCCGGCTCAGCTAAGCTCCAACTA
CCTCCTGTCTTGCAGATCCCGTACAAGGTGGTGGCACGGCGGGAAGGTGATGTGGCAGCCTGTTACGCCA
ACCCCAGCCTGGCCCAAGAGGAGCTGGGGTGGACAGCAGCCTTAGGGCTGGACAGGATGTGTGAGGATCT
CTGGCGCTGGCAGAAGCAGAATCCTTCAGGCTTTGGCACGCAAGCCTGAGGACCCTCCCCTACCAAGGAC
CAGGAAAAGCAGCAGCTGCCTGCTCTCCAGCCTCTGGCAGGAACTCAGGGCCCTGGAGCTGCTGGGGCCA
AGCCAAGGCCTCCCCTACCTCAAACCCCAGCTGGGCCCGCTTAGCCCACCAGGCATGAGGCCAAGGGCTC
CACTGACCAGGAGGCCGAGGTCTCTAACTCTTATCTTCCACAGGGTCCAAGAGTTCATCAGGACCCCCAA
GAGTGAGTGAGGGGGCAAGGCTCTGGCACAAAACCTCCTCCTCCCAGGCACTCATTTATATTGCTCTGAA
AGAGCTTTCCAAAGTATTTAAAAAATAAAAACAAGTTTTCTTACACTGGG

>DEX0367_024.nt.4
CGTGGAGGGGCGGGGTCGGGGCCGGAGTGGGCCGTCAGCACTTAAAGGGCCCGCGGCTCGGGCGTAGGAG
GCGGTGCCTCTGCAGCAAGCGTGGGGCGCGGGAACCCGAGCAGGACTCTCCAGTCCTCAGTCACCTTGGA
CAAAGAAGTGTGGATCCTCAGATTCCATCTTTTCCAACTCCAAGGTGCCATGGCAGAGAAGGTGCTGGTA
ACAGGTGGGGCTGGCTACATTGGCAGCCACACGGTGCTGGAGCTGCTGGAGGCTGGCTACTTGCCTGTGG
TCATCGATAACTTCCATAATGCCTTCCGTGGAGGGGGCTCCCTGCCTGAGAGCCTGCGGCGGGTCCAGGA
GCTGACAGGCCGCTCTGTGGAGTTTGAGGAGATGGACATTTTGGACCAGGGAGCCCTACAGCGTCTCTTC
AAAAAGTACAGCTTTATGGCGGTCATCCACTTTGCGGGGCTCAAGGCCGTGGGCGAGTCGGTGCAGAAGC
CTCTGGATTATTACAGAGTTAACCTGACCGGGACCATCCAGCTTCTGGAGATCATGAAGGCCCACGGGGT
GAAGAACCTGGTGTTCAGCAGCTCAGCCACTGTGTACGGGAACCCCCAGTACCTGCCCCTTGATGAGGCC
CACCCCACGGGTGGTTGTACCAACCCTTACGGCAAGTCCAAGTTCTTCATCGAGGAAATGATCCGGGACC
TGTGCCAGGCAGACAAGACTTGGAACGCAGTGCTGCTGCGCTATTTCAACCCCACAGGTGCCCATGCCTC
TGGCTGCATTGGTGAGGATCCCCAGGGCATACCCAACAACCTCATGCCTTATGTCTCCCAGGTGGCGATC
GGGCGACGGGAGGCCCTGAATGTCTTTGGCAATGACTATGACACAGAGGATGGCACAGGTGTCCGGGATT
ACATCCATGTCGTGGATCTGGCCAAGGGCCACATTGCAGCCTTAAGGAAGCTGAAAGAACAGTGTGGCTG
CCGGGTAGGAAGAGAAGGGAGAAGTGAGGGAGGGGAGGGACCAGATCCAGGCAGAGCGGCCCAGAGGCGT
GGTCAGTCCAGCCCGCTCCATAAACCCTGTTCTCCGTGGGCCAGATCTACAACCTGGGCACGGGCACAGG

CTATTCAGTGCTGCAGATGGTCCAGGCTATGGAGAAGGCCTCTGGGAAGAAGGTAGGCCCCCCCACCCCA
CCCACCTCACCCACCTCACCCCACCCAGCTCTCAGCAACAGAGCAGCTTTGGCCCTTCCCACTCACCTCT
CTGGAGGGTACCTGGCACTGCCCAGTCTCCTCAGCCTGCCTAGCACCCGGTGCCTGCGGGCCAGCCGCTG
CTCTGCACTTCTGCTTCTTAAAGACCTGGCCAGCAGCTGGGCCAGGGCCGGCTCAGCTAAGCTCCAACTA

5   CCTCCTGTCTTGCAGATCCCGTACAAGGTGGTGGCACGGCGGGAAGGTGATGTGGCAGCCTGTTACGCCA
ACCCCAGCCTGGCCCAAGAGGAGCTGGGGTGGACAGCAGCCTTAGGGCTGGACAGGATGTGTGAGTGCTG
GCCTGACCCCTGCAGTTAGGGGCCAGGGAGGGGCTAAATCAGATCTGGGCATGCCCACCCAGGGCGGGCT
GTGGGGTGGGGGAGGCTCTGCATTCCACCGTGGGCCCTGAGCACACACCTCGCCGTGTTGCAGGTGAGGA
TCTCTGGCGCTGGCAGAAGCAGAATCCTTCAGGCTTTGGCACGCAAGCCTGAGGACCCTCCCCTACCAAG

10 GACCAGGAAAAGCAGCAGCTGCCTGCTCTCCAGCCTCTGGCAGGAACTCAGGGCCCTGGAGCTGCTGGGG
CCAAGCCAAGGCCTCCCCTACCTCAAACCCCAGCTGGGCCCGCTTAGCCCACCAGGCATGAGGCCAAGGG
CTCCACTGACCAGGAGGCCGAGGTCTCTAACTCTTATCTTCCACAGGGTCCAAGAGTTCATCAGGACCCC
CAAGAGTGAGTGAGGGGGCAAGGCTCTGGCACAAAACCTCCTCCTCCCAGGCACTCATTTATATTGCTCT
GAAAGAGCTTTCCAAAGTATTTAAAAATAAAAACAAGTTTTCTTACACTGGG

15

>DEX0367_025.nt.1
CTATTGGTCAGGACGTTTCCTATGCTAATAAAGGGGTGGCCCGTAGAAGATTCCAGCACCCTCCCCTAAC
TCCAGGCCAGACTCCTTTCAGCTAAAGGGGAGATCTGGATGGCATCTACTTCGTATGACTATTGCAGAGT
GCCCATGGAAGACGGGGATAAGCGCTGTAAGCTTCTGCTGGGGATAGGAATTCTGGTGCTCCTGATCATC

20 GTGATTCTGGGGGTGCCCTTGATTATCTTCACCATCAAGGCCAACAGCGAGGCCTGCCGGGACGGCCTTC
GGGCAGTGATGGAGTGTCGCAATGTCACCCATCTCCTGCAACAAGAGCTGACCGAGGCCCAGAAGGGCTT
TCAGGATGTGGAGGCCCAGGCCGCCACCTGCAACCACACTGTGATGGCCCTAATGGCTTCCCTGGATGCA
GAGAAGGCCCAAGGACAAAAGAAAGTGGAGGAGCTTGAGGGAGAGATCACTACATTAAACCATAAGCTTC
AGGACGCGTCTGCAGAGGTGGAGCGACTGAGAAGAGAAAACCAGGTCTTAAGCGTGAGAATCGCGGACAA

25 GAAGTACTACCCCAGCTCCCAGGACTCCAGCTCCGCTGCGGCGCCCCAGCAGCACAATCAGCAGCTGGGG
CGCCGCAGCTGCTGATTGTGCTGCTGGGCCTCAGCGCTCTGCTGCAGTGAGATCCCAGGAAGCTGGCACA
TCTTGGAAGGTCCGTCCTGCTCGGCTTTTCGCTTGAACATTCCCTTGATCTCATCAGTTCTGAGCGGGTC
ATGGGGCAACACGGTTAGCGGGGAGAGCACGGGGTAGCCGGAGAAGGGCCTCTGGAGCAGGTCTGGAGGG
GCCATGGGGCAGTCCTGGGTGTGGGGACACAGTCGGGTTGACCCAGGGCTGTCTCCCTCCAGAGCCTCCC

30 TCCGGACAATGAGTCCCCCCTCTTGTCTCCCACCCTGAGATTGGGCATGGGGTGCGGTGTGGGGGGCATG
TGCTGCCTGTTGTTATGGGTTTTTTTTGCGGGGGGGGTTGCTTTTTTCTGGGGTCTTTGAGCTCCAAAAA
ATAAACACTTCCTTTGAGGGAGAGCACACCT

>DEX0367_025.nt.2
35 CTATTGGTCAGGACGTTTCCTATGCTAATAAAGGGGTGGCCCGTAGAAGATTCCAGCACCCTCCCCTAAC
TCCAGGCCAGACTCCTTTCAGCTAAAGGGGAGATCTGGATGGCATCTACTTCGTATGACTATTGCAGAGT
GCCCATGGAAGACGGGGATAAGCGCTGTAAGCTTCTGCTGGGGATAGGAATTCTGGTGCTCCTGATCATC
GTGATTCTGGGGGGTGCCCTTGATTATCTTCACCATCAAGGCCAACAGCGAGGCCTGCCGGGACGGCCTTC

40 GGGCAGTGATGGAGTGTCGCAATGTCACCCATCTCCTGCAACAAGAGCTGACCGAGGCCCAGAAGGGCTT
TCAGGATGTGGAGGCCCAGGCCGCCACCTGCAACCACACTGTGATGGCCCTAATGGCTTCCCTGGATGCA
GAGAAGGCCCAAGGACAAAAGAAAGTGGAGGAGCTTGAGGTCACTACATTAAACCATAAGCTTCAGGACG
CTTAAGACCTGGTTTTCTCTTCTCAGTCGCTCCACCTCTGCAGACGCGTCTGAGAAGAGAAAACCAGGTC
TTAAGCGTGAGAATCGCGGACAAGAAGTACTACCCCAGCTCCCAGGACTCCAGCTCCGCTGCGGCGCCCC

45 AGCTGCTGATTGTGCTGCTGGGCCTCAGCGCTCTGCTGCAGTGAGATCCCAGGAAGCTGGCACATCTTGG
AAGGTCCGTCCTGCTCGGCTTTTCGCTTGAACATTCCCTTGATCTCATCAGTTCTGAGCGGGTCATGGGG
CAACACGGTTAGCGGGGAGAGCACGGGGTAGCCGGAGAAGGGCCTCTGGAGCAGGTCTGGAGGGGCCATG
GGGCAGTCCTGGGTGTGGGGACACAGTCGGGTTGACCCAGGGCTGTCTCCCTCCAGAGCCTCCCTCCGGA
CAATGAGTCCCCCCTCTTGTCTCCCACCCTGAGATTGGGCATGGGGTGCGGTGTGGGGGGCATGTGCTGC

50 CTGTTGTTATGGGTTTTTTTTGCGGGGGGGGTTGCTTTTTTCTGGGGTCTTTGAGCTCCAAAAAATAAAC
ACTTCCTTTGAGGGAGAGCACACCT

>DEX0367_025.nt.3
CTCCGAAGTCTAGAGAGAGAGAGGCTGCAGGGGAAGGGACTGGGCGCAGAGGTGAGGGGCCTTGGGGAAGGA
GGGAGAATCCAGTGAGAAGGATGCCCTGGATCGGGGGCGGGCTGATAGCTGTATGGGGCGGGGCTATCAA

55 TGGGGGCGGGTCTATCGATGGAGGGGAGGGGCTTATGGATGGGTGGGGCGGGTCTAGCGATGGGGGCGGG
GCTTGGGGCCGGGCTTGGAGCGGGGCCAGTGCCTGCTGCCTCCAGTCTGCCCTGAGTCCCTCTTCTGGTC
CTTTAGGCACATCTTGGAAGGTCCGTCCTGCTCGGCTTTTCGCTTGAACATTCCCTTGATCTCATCAGTT
CTGAGCGGGTCATGGGGCAACACGGTTAGCGGGGAGAGCACGGGGTAGCCGGAGAAGGGCCTCTGGAGCA
GGTCTGGAGGGGCCATGGGGCAGTCCTGGGTGTGGGGACACAGTCGGGTTGACCCAGGGCTGTCTCCCTC

60 CAGAGCCTCCCTCCGGACAATGAGTCCCCCCTCTTGTCTCCCACCCTGAGATTGGGCATGGGGTGCGGTG

TGGGGGGCATGTGCTGCCTGTTGTTATGGGTTTTTTTTGCGGGGGGGGGTTGCTTTTTTTCTGGGGTCTTTG
AGCTCCAAAAAATAAACACTTCCTTTGAGGGAGAGCACACCT

>DEX0367_026.nt.1

5 GATCTCTTCCAAATGTCCCCGCTCTCCCCAGGCTCTCCCCTCCCGCCACTTGCCAGGGCTGACCTCACCG
CCATCTTAACCGGGTGTCCACCTCTCTCTGCCTGCCTGGTGCTGGCCCCGCGTCCCCATCGCCGCGCCCG
TCTGCTCCCCTCAGAGGGCTTGCTGACGCTGCGGGCCAAGCCGCCCTCGGAGGCCGAGTACACCGACGTG
CTGCAGAAGATCAAGTACGCCTTCAGCCTGCTGGCCCGGCTGCGCGGCAACATCGCCGACCCCTCCTCTC
CGGAGCTGTTGCACTTCCTTTTCGGGCCTCTGCAGATGATTGTGAACACGTCGGGGGGGCCGGAGTTCGC
10 GAGCAGTGTGCGGCGGCCGCATCTGACATCGGATGCCGTGGCGCTGCTGCGGGCAACGTCACTCCACGT
GAAAACGAGCTCTGGACCTCGCTGGGGGACTCGTGGACCCGCCCCGGGCTGGAGCTGTCCCCGGAGGAGG
GACCCCCATACAGACCCGAGTTCTTCAGCGGCTGGGAGCCGCCGGTCACTGACCCGCAGAGCCGCGCCTG
GGAGGACCCAGTTGAGAAACAGCTACAGCACGAGCGGAGGCGGCAGCAAAGCGCCCCCCAGGTCGCT
GTCAATGGTCACCGAGACTTGGAGCCAGAATCTGAGCCTCAGCTGGAGTCAGAGACAGCAGGAAAATGGG
15 TCCTGTGTAATTATGACTTCCAGGCCCGCAACAGCAGTGAGCTGTCGGTCAAGCAGCGGGACGTACTGGA
GGTCCTGGATGACAGTCGTAAGTGGTGGAAGGTTCGGGACCCAGCGGGGCAGGAGGGATATGTGCCCTAC
AACATCCTGACACCCTACCCCGGACCCCGGCTGCACCACAGCCAAAGCCCTGCCCGCAGCCTGAACAGCA
CTCCTCCTCCACCACCAGCCCCAGCCCCGGCCCCACCTCCAGCTCTGGCTCGGCCCCGCTGGGACAGGCC
CCGCTGGGACAGCTGCGATAGCCTCAACGGCTTGGACCCCAGCGAGAAGGAGAAATTCTCCCAGATGCTC
20 ATCGTCAACGAGGAACTGCAGGCGCGCCTGGCCCAGGGCCGCTCGGGACCGAGCCGCGCAGTCCCAGGGC
CCCGCGCCCCGGAACCGCAGCTCAGCCCGGGCTCGGACGCCTCCGAGGTCCGCGCCTGGCTGCAGGCCAA
GGGCTTTAGCTCCGGGACCGTGGACGCGCTGGGTGTGCTGACCGGGGCGCAGCTTTTCTCGCTGCAGAAG
GAGGAGCTGCGGGCGGTGAGCCCCGAGGAGGGGGGCACGTGTGTACAGCCAGGTCACCGTGCAGCGCTCGC
TGCTGGAGGACAAAGAGAAAGTGTCAGAGCTGGAGGCAGTGATGGAGAAGCAAAAGAAGAAGGTGGAAGG
25 CGAGGTGGAAATGGAGGTCATTTGACCTGCCAGGCGCCCTTCGCAAAGAGTGACGAGGCCCCGTGGGAGA
ACGGACTCCTCAGACTCTCCCCAATAGCGGAAGTCGATCTTCTGAAGGATGGCCAATCTGCTCCGGCCCT
GGTCTTCCCCCATCCCGGTGGACAGACTTAACGATCCTTGCTGCAGTCCCTCCGGAGAGGATCTGGACTG
GCTGGGAGTGGGGGAGGGCGTGGAGACAGTCTACGAAAGCGCTAGCAGACCCCCGAGAGGGTGCAGTGGA
GCCCTGAGCATTGTAATATGCGGCCCAGCCTATAAACAGCCTCCGTGCTTAGCAGATGGTGTGCCAACTG
30 CTGTCCTCTACCCGGCTACCAGGGGGCTCAGAAACCTGCCCCTCGACCCAGTCACAGCCTCATGGAACCT
G

>DEX0367_027.nt.1
AGGCAGACGGGAACCTGAGACTTTGTCCAAACCTCGGCTGGAAGGCTGGTCTTGGCACTCGGCGTCTTAG
35 CTAGTGGATTTATCGCATTAAGACTGTACTTGGGCCTTTTTCATGATGACTTGGGTCTCAATCCTTGGAG
CAGCTGAAAGAACAGCTCGGGCCACCTGTGGATGCACCCACAATGAAATCAGAGGCCACGTGAGGTTTGC
AAGGCAGAGTGGCCTGCAGTTTTAACTCGGGCACTATGGCAATGCAGACCCAAACCAATTGTGCCAGTTC
TTTGCTGGGTGGCAAAGTGTTTTTCAGGCTTGTGGTCACTCCTGATTTCTGTTTGGAAGCACAGTTCTCTC
TCTAAGGAGAATGGAAACTTGGAAAGGCCTGTAAAATCGGAAAGCCTGTAAAAGGCCGTAGAGTGGAGAT
40 GTATAAAAATGTACGGTCACCCATGTTTTAGGGGTTGACCGGGAGCGGGGCGAGGAGGAGGGTGGGTATG
AGGCGGTAGCGGACCGCGGAGTCGAGACCTACCCGAACGACGCGGGCGAGCGGGGCTTTGGACGCCGGTG
GAGACGCAGGCGAGGGCTGCTGTAAAATATTACAGGCATACTTTGGAGATGTTATGGGTTTGGTTCCACA
TCACTGCAATAAAACTCCAGCGGGAGGAAGAAGAGGCCTTTGCCAGCAGTCAGAGCAGCCAAGGGGCCCA
ATCCCTCATATTCTCCAAGTTTGAAGGAAAGAAAACCAACAAGAAGACCCGCAAGGTTACCACAGTGAAG
45 AAATCTTCAGTACGTCTTCCAGGGTCGGATCAAAGAAGGATATTGAAATGGATTCCAGGCGTGTGCCTCG
AGACAAGCTGGCCTGCATCACCAAGTGCAGCAAGCACATCTTCGATGCCATCAAGATCACCTAGAACGAG
CTGGCGTCAGCAGATGACTTCCTCCCCACCCTCATCTACATTGTTTTGAAGGGCAACCCCCATGCCTTCA
GTCTAATATCCAGTATATCACGCGCTTCTGCAATCCAAGCCGACTGATGACTGGAGAGGATGGCTACTAT
TTCACCAATCTGGTGAGTAAGTGAGTTCTTGGCGTTGTGGAGAAGGACTAGGAAGGTGGTGGTTTTGGGG
50 ATGTGATAGGTCACTCAGGCCCATGACAGGTGAATGCTTCTGTGTGAGAAGGCAGCACGGCTGAGGAAGC
TCACTTTGCATCAGGGAGCACAAGGACCAGGCCGTACAGACACTCCGCCTCCCCGCACTTGATCAGAGAT
TGTGTTTATCCTACAGAAACAGATGACATGTGTTGGGCATCACTCCCACGGTCCTGGGTAGAAGAGTCC
TTCACTTGGCAGGGCTTTTTCAACCAATGAATAAGGCAAATTATATATAAGTTAATAATGCCATTTCGAA
CCGAGACAGATGGCAGCTAAATGAAGTTTAATTAAAGAATGAGCGCGGGCCCTTTTTCATTGGGTACTGC
55 ATCTACTTCGACCGCAAAAGACGAAGTGACCCCAACTTCAAGAACAGGCTTCGAGAACGAAGAAAGAAAC
AGAAGCTTGCCAAGGAGAGAGCTGGGCTTTCCAAGTTACCTGACCTTAAAGATGCTGAAGCTGTTCAGAA
GTTCTTCCTTGAAGAAATACAGCTTGGTGAAGAGTTACTAGCTCAAGGTGAATATGAGAAGGGCGTAGAC
CATCTGACAAATGGAATTGCTGTGTGTGGACAGCCACAGCAGTTACTGCAGGAATTCTAAGTGCTCAGAGCTTG
60 GCTGAAGATGATGTGGAATGAGAAACAAATGTTAACATAATAAAATCTCAGTTAAAAAATATTTAAAAAAT

TCTTGGTAGTTGAGCAGCTCTGGGGGAATAAGGGCAAATATCCTTGTTATGAACTACACTGAAATCTACC
AAAGTTAATGTTTACTTTGTGTAGGTCCATTTGTCTATTTTATTTATTTTTCCCAGTGAAAAGTGTATTT
TGATAGAGAACTTTTCATTTTATAAATACACTATGAGTTACTGAAAATTCATATCATGGATTTCATTTAT
TCCTGAAACATAGTTAAAATGTACATATGACATGGCTTATGTTAAAAATACCCAGTGCTCAGTTTTGAAA

5   GATAGGCACAAAAA

>DEX0367_028.nt.1
CTGGCGAGGTGGCCCGGGCTCGGCCCTCCGAGGGTGTGGAGCGCCCGGGAAACGGGACCCCAAGAATCCA
ACTGCTCCCACCCTGCGCGGCCAGGAGCGGGGCTGGCGATACTGGGGACACTGTCCAATCCGGGCAGGGG

10 GCGAGGCCGGGGGACCGGGCGGGCCTGGGGAGCAGGCACGTGTGGAGGGCGGACCCGCCGGGGGGTCGAGG
CCTGCCTCTCCGAGAGCTCCTGGCGCGGCCGTCCCGGCCCGGGGCCCCAGGTGCGCTTCCCCTAGAGAGG
GATTTTCCGGTCTCGTGGGCAGAGGAACAACCAGGAACTTGGGCTCAGTCTCCACCCCACAGTGGGGCGG
ATCCGTCCCGGATAAGACCCGCTGTCTGGCCCTGAGTAGGGTGTGACCTCCGCAGCCGCAGAGGAGGAGC
GCAGCCCGGCCTCGAAGAACTTCTGCTTGGGTGGCTGAACTCTGATCTTGACCTAGAGTCATGGCCATGG

15 CAACCAAAGGAGGTACTGTCAAAGCTGCTTCAGGATTCAATGCCATGGAAGATGCCCAGACCCTGAGGAA
GGCCATGAAAGGGCTCGCACAAACAACGAGGAAACAGAAGATTCACCAAACATTTATGAACAGCCAGCTG
TTATAAAGACAAGGGACCAATCTGAAAAATCAGAATAAAGTCTCCTTGTTAAAGCTGCTGCAAGAAACAG
GCACCGATGAAGACGCCATTATTAGCGTCCTTGCCTACCGCAACACCGCCCAGCGCCAGGAGATCAGGAC
AGCCTACAAGAGCACCATCGGCAGGGACTTGATAGACGACCTGAAGTCAGAACTGAGTGGCAACTTCGAG

20 CAGGTGATTGTGGGGATGATGACGCCCACGGTGCTGTATGACGTGCAAGAGCTGCGAAGGGCCATGAAGG
GAGCCGGCACTGATGAGGGCTGCCTAATTGAGATCCTGGCCTCCCGGACCCCTGAGGAGATCCGGCGCAT
AAGCCAAACCTACCAGCAGCAATATGGACGGAGCCTTGAAGATGACATTCGCTCTGACACATCGTTCATG
TTCCAGCGAGTGCTGGTGTCTCTGTCAGCTGGTGGGAGGGATGAAGGAAATTATCTGGACGATGCTCTCG
TGAGACAGGATGCCCAGGACCTGTATGAGGCTGGAGAGAAGAAATGGGGGACAGATGAGGTGAAATTTCT

25 AACTGTTCTCTGTTCCCGGAACCGAAATCACCTGTTGCATGTGTTTGATGAATACAAAAGGATATCACAG
AAGGATATTGAACAGAGTATTAAATCTGAAACATCTGGTAGCTTTGAAGATGCTCTGCTGGCTATAGTAA
AGTGCATGATGAGGAACAAATCTGCATATTTTGCTGAAAAGCTCTATAAATCGATGAAGGGCTTGGGCACCGA
TGATAACACCCTCATCAGAGTGATGGTTTCTCGAGCAGAAATTGACATGTTGGATATCCGGGCACACTTC
AAGAGACTCTATGGAAAGTCTCTGTACTCGTTCATCAAGGGTGACACATCTGGAGACTACAGGAAAGTAC

30 TGCTTGTTCTCTGTGGAGGAGATGATTAAAAATAAAAATCCCAGAAGGACAGGAGGATTCTCAACACTTTG
AATTTTTTTAACTTCATTTTTCTACACTGCTATTATCATTATCTCAGAATGCTTATTTCCAATTAAAACG
CCTACAGCTGCCTCCTAGAATATAGACTGTCTGTATTATTATTCACCTATAATTAGTCATTATGATGCTT
TAAAGCTGTACTTGCATTTCAAAGCTTATAAGATATAAATGGAGATTTTAAAGTAGAAATAAATATGTAT
TCCATGTTTTTAAAAGATTACTTTCTACTTTGTGTTTCACAGACATTGAATATATTAAATTATTCCATAT

35 TTTCTTTTCAGTGAAAAATTTTTTAAATGGAAGACTGTTCTAAAATCACTTTTTTTCCCTAATCCAATTTTT
TAGAGTGGCTAGTAGTTCTTCATTTGAAATTGTAAGCATCCGGTCAGTAAGAATGCCCATCCAGTTTTTC
TATATTTCATAGTCAAAGCCTTGAAAGCATCTACAAATCTCTTTTTTTTAGGTTTTTGTCCATAGCATCAGT
TGATCCTTACTAAGTTTTTTCATGGGAGACTTCCTTCATCACATCTTATGTTGAAATCACTTTCTGTAGTC
AAAGTATACCAAAACCAATTTATCTGAACTAAATTCTAAAGTATGGTTATACAAACCATATACATCTGGT

40 TACCAAACATAAATGCTGAACATTCCATATTATTATAGTTAATGTCTTAATCCAGCTTGCAAGTGAATGG
AAAAAAAAATAAGCTTCAAACTAGGTATTCTGGGAATGATGTAATGCTCTGAATTTAGTATGATATAAAG
AAAACTTTTTTGTGCTAAAAATACTTTTTAAAATCAATTTTGTTGATTGTAGTAATTTCTATTTGCACTG
TGCCTTTCAACTCCAGAAACATTCTGAAGATGTACTTGGATTTAATTAAAAAGTTCACTTTGTAA

45 >DEX0367_028.nt.2
AAGATTCAAGGTCTTGCTATGTTTCCCAGGCCAGCTCTTCACAGGTGCCATCTTAGCTCACTGCCTTGAA
CTCCTGGCCTCAAGACATCCTCCCACCTCAGCCTCGAGTAGCTGGAACTACATGCACATCCCGACCCCCT
TTTGGCAGAGGTTGATCAATGTAGATTTGCCAGTTAGGAAGGCATGAGTCAGCTGGGAAGACACAGGAGA
GTGGGGTGGTGGGGGAGGCCAGGAGATAGTGGAGGGAGCTGTCCTTGGGAAGCTTCTCCTGCTATTGGCA

50 GAGGCACTCATGCAACGTTAGGGGCCACACTGAGGCTCAGCCCTCTAGGCACACATCAGAGCCATTGTAA
TAATGCGAGCGCATTTAGAGGACGTCTGTCCTCCCTCCCTGGGTCTGAAGCTCCAATTCTTGTGGTATTT
TGTCTCATTACTCTAGACCTGCTATCTTAATGGTGTCCCAGATCTTCATGTATACTCTCGAATTATTTT
TTATTTGTTTTTCTCATCCTCAGGGTGGGAGGGATGAAGGAAATTATCTGGACGATGCTCTCGTGAGACA
GGATGCCCAGGACCTGTATGAGGCTGGAGAGAAGAAATGGGGGACAGATGAGGTGAAATTTCTAACTGTT

55 CTCTGTTCCCGGAACCGAAATCACCTGTTGCATGTGTTTGATGAATACAAAAGGATATCACAGAAGGATA
TTGAACAGAGTATTAAATCTGAAACATCTGGTAGCTTTGAAGATGCTCTGCTGGCTATAGTAAAGTGCAT
GAGGAACAAATCTGCATATTTTGCTGAAAAGCTCTATAAATCGATGAAGGGCTTGGGCACCGATGATAAC
ACCCTCATCAGAGTGATGGTTTCTCGAGCAGAAATTGACATGTTGGATATCCGGGCACACTTCAAGAGAC
TCTATGGAAAGTCTCTGTACTCGTTCATCAAGGGTGACACATCTGGAGACTACAGGAAAGTACTGCTTGT

60 TCTCTGTGGAGGAGATGATTAAAAATAAAAATCCCAGAAGGACAGGAGGATTCTCAACACTTTGAATTTTT

```
TTAACTTCATTTTTTCTACACTGCTATTATCATTATCTCAGAATGCTTATTTCCAATTAAAACGCCTACAG
CTGCCTCCTAGAATATAGACTGTCTGTATTATTATTCACCTATAATTAGTCATTATGATGCTTTAAAGCT
GTACTTGCATTTCAAAGCTTATAAGATATAAATGGAGATTTTTAAAGTAGAAATAAATATGTATTCCATGT
TTTTAAAAGATTACTTTCTACTTTGTGTTTCACAGACATTGAATATATTAAATTATTCCATATTTTCTTT
```
5
```
TCAGTGAAAAATTTTTTAAATGGAAGACTGTTCTAAAATCACTTTTTTCCCTAATCCAATTTTTAGAGTG
GCTAGTAGTTTCTTCATTTGAAATTGTAAGCATCCGGTCAGTAAGAATGCCCATCCAGTTTTCTATATTT
CATAGTCAAAGCCTTGAAAGCATCTACAAATCTCTTTTTTTAGGTTTTGTCCATAGCATCAGTTGATCCT
TACTAAGTTTTTCATGGGAGACTTCCTTCATCACATCTTATGTTGAAATCACTTTCTGTAGTCAAAGTAT
ACCAAAACCAATTTATCTGAACTAAATTCTAAAGTATGGTTATACAAACCATATACATCTGGTTACCAAA
```
10
```
CATAAATGCTGAACATTCCATATTATTATAGTTAATGTCTTAATCCAGCTTGCAAGTGAATGGAAAAAAA
AATAAGCTTCAAACTAGGTATTCTGGGAATGATGTAATGCTCTGAATTTAGTATGATATAAAGAAAACTT
TTTTGTGCTAAAAATACTTTTTAAAATCAATTTTGTTGATTGTAGTAATTTCTATTTGCACTGTGCCTTT
CAACTCCAGAAACATTCTGAAGATGTACTTGGATTTAATTAAAAAGTTCACTTTGTAA
```

15 >DEX0367_029.nt.1
```
GAAAAAAAAAAAGAAAAAAAACAAAAAACAAAAAATACATCACAATAAGAGGCCGAATGTAAATAAAGAC
GAGGGACGCAGAGGGGTGGCCGTGGCTGAGAGGAGACAGCGCCGCAGCACTGAGGGTTTGGGCTTGCAGG
CGCTGCAGGAGACGCCCAGGCGGAGTCTTGTCTCGCAGCCAGCTCTGAGCGGGAGGCCTGAGCGGGAAGC
ATTGGCGTCCGAGCGACTTCTAGGAGCCTGGGGTTCGGCGCTATGGAGGAGCTCGATGGCGAGCCAACAG
```
20
```
TCACTGTAAGGGTACCCCGAACAGGCTTGCTCGTCCTTGCGGGTTGAGAACTGCGTCTGCTTAGTTACTT
CAGGCTTGTCTGCTTCCCTAGTGGTCGCGAGGCGCTCGTCCCCTTCCCTCGACTCAGTTGCCACTTTTCC
GGAGGTCGCAGTGTTAACGAGTTGATTCCAGGCGTGAATTCCAAGAAGAACCAAATGTATTTTGACTGGG
GTCCAGGGGAGATGCTGGTATGTGAAACCTCCTTCAACAAAAAAGAAAAATCAGAGATGGTGCCAAGTTG
CCCCTTTATCTATATCATCCGTAAGGATGTAGATGTTTACTCTCAAATCTTGAGAAAACTCTTCAATGAA
```
25
```
TCCCATGGAATCTTTCTGGGCCTCCAGAGAATTGACGAAGAGTTGACTGGAAAATCCAGAAAATCTCAAT
TGGTTCGAGTGAGTAAAAACTACCGATCAGTCATCAGAGCATGTATGGAGGAAATGCACCAGGTTGCAAT
TGCTGCTAAAGATCCAGCCAATGGCCGCCAGTTCAGCAGCCAGGTCTCCATTTTGTCAGCAATGGAGCTC
ATCTGGAACCTGTGTGAGATTCTTTTTATTGAAGTGGCCCCAGCTGGCCCTCTCCTCCTCCATCTCCTTG
ACTGGGTCCGGCTCCATGTGTGCGAGGTGGACAGTTTGTCGGCAGATGTTCTGGGCAGTGAGAATCCAAG
```
30
```
CAAACATGACAGCTTCTGGAACTTGGTGACCATCTTGGTGCTGCAGGGCCGGCTGGATGAGGCCCGACAG
ATGCTCTCCAAGGAAGCCGATGCCAGCCCCGCCTCTGCAGGCATATGCCGAATCATGGGGGACCTGATGA
GGACAATGCCCATTCTTAGTCCTGGGAACACCCAGACACTGACAGAGCTGGAGCTGAAGTGGCAGCACTG
GCACGAGGAATGTGAGCGGTACCTCCAGGACAGCACATTCGCCACCAGCCCTCACCTGGAGTCTCTCTTG
AAGATTATGCTGGGAGACGAAGCTGCCTTGTTAGAGCAGAAGGAACTTCTGAGTAATTGGTATCATTTCC
```
35
```
TAGTGACTCGGCTCTTGTACTCCAATCCCACAGTAAAACCCATTGATCTGCACTACTATGCCCAGTCCAG
CCTGGACCTGTGTTTCTGGGAGGTGAGAGCAGCCCAGAACCCCTGGCAACATCTTGTTGGCAGCCTTTGAG
TTTGACATCCATCAAGTAATCAAAGAGTGCAGGAATAAAAACAGATTTGAGCAGAAGAAGTCTGTTGGATG
CAGGGAGTATAAAAGGGGAAAGCATCCTGCTGTTTCCAGTGGCAGAAGAGAAAGAGAAGTACCATGAAGA
GGGGTGAACAACATCACTCTCGTGCCGGACACAGACAAGGAATGATTAGGCTTGCAGTCCCCCAGTCTCA
```
40
```
GCCTCTCTCCACTCTCACAAGCAGGAAGGGATCTGAGCTCTGTTAGTGGGAAATCTTGGAGTTTGCATTT
CTCTGATCACGAGGCTACCTCCTATCCTGTTTTGAAAGAGGAACAGAAAAGCAGAGCTTGTCCCAGGAAT
TGTGAAATGGGCTGCAATGTCTTCAGAGGCAGTAGCAGTTTTAGGAAGTGTGCCCTCCTGCCTCTGGGGA
AAGGTTGAATTGGGGTCTTGTGGGCACTTGTTTTCTCATGGGATGTATCAGTCAGACCCATGCACAGTTC
TGGGCGGCAGGGAAGTGTGTCTCTTTCTGGACTTGAGCCAGTGAGCTGTTGGCATGCCTTTTTAAAGGCA
```
45
```
TTAGAAAAAAGGTTTCTTTCTTTTTTTTTCCTTTCCTGTTCACACCCACCTTCATATGAACATGGAGAAGGG
AGGTTTCTAAGTGGGCCCCTGATTTACTCCTATGTCAATAATCGTCCAGACCCTAAGGGGCAGAAAAGG
CAGCCAGGGGATCCTGACTCAGAGGCTGTAGTTCGTCAGCCCTGCTAAAAGGGCCAGGGGTTTGTCTCTCA
GTGCCTGGGTGGGCATGCTTCTTGACTGCTAACCCCTTAGCTGAAAGGAGATGGCAGGAGTCCTCCACAGGGT
CCCTTCTTCCCTCCTGGCTTACAGCTTCCTCTGAAGACGGGCCAGTTCCTTGCTGTCTGGCAAGAGGAAA
```
50
```
CAGCTGGGGTGCACTTCACTGGGAGTTGGGCAAGATGTCGGCAGTTTCCTGGTGCCCTTCAAGTCCTCCA
AAAGTATCGAGCTAAAAGAGGTGACAAGTCTTCCTTCAATAACCAAGAAGAGTTCCAGCTTAATGTCTGC
TGAATCTAGAATCTTCTGACTTGTATTTTGGCCCAGCCCACAATCTTGGATGCAGGGTGCAGCAGTGTTT
ACAAACCTCACACCCCATTTCCAGAATAGCAGGTAGAGGCACTGGAGCTTTGATTACTCCAGAGCCCCCA
ACAAGCAAGATGGGGCCTGGCATCGCCCTGAGCAACTGGTGGTTTGTGGCCCACCTGACAGACCTGCTGG
```
55
```
ACCACTGCAAGCTCCTCCAGTCACACAACCTCTATTTCGGTTCCAACATGAGAGAGTTCCTCCTGCTGGA
GTACGCCTCGGGACTGTTTGCTCATCCCAGCCTGTGGCAGCTGGGGGTCGATTACTTTGATTACTGCCCC
GAGCTGGGCCGAGTCTCCCTGGAGCTGCACATTGAGCGGATACCTCTGAACACCGAGCAGAAAGCCCTGA
AGGTGCTGCGGATCTGTGAGCAGCGGCAGATGACTGAACAAGTTCGCAGCATCCGTGTGTAAGGATGCCGCCTTT
```
60 `GCCACGCTCGTGTCAGACAGGTTCCTCAGGGATTACTGTGAGCGAGGCTGCTTTTCTGATTTGGATCTCA`

TTGACAACCTGGGGCCAGCCATGATGCTCAGTGACCGACTGACATTCCTGGGAAAGTATCGCGAGTTCCA
CCGTATGTACGGGGAGAAGCGTTTTGCCGACGCAGCTTCTCTCCTTCTGTCCTTGATGACGTCTCGGATT
GCCCCTCGGTCTTTCTGGATGACTCTGCTGACAGATGCCTTGCCCCTTTTGGAACAGAAACAGGTGATTT
TCTCAGCAGAACAGACTTATGAGTTGATGCGGTGTCTGGAGGACTTGACGTCAAGAAGACCTGTGCATGG
AGAATCTGATACCGAGCAGCTCCAGGATGATGACATAGAGACCACCAAGGTGGAAATGCTGAGACTTTCT
CTGGCACGAAATCTTGCTCGGGCAATTATAAGAGAAGGCTCACTGGAAGGTTCCTGAGAACTGCTTCAAT
GTGGTATCTTTGTATGGCAATGTATATAGATTTTTTAAAAGAATAAATGTTGTTTTGCAAATG

>DEX0367_029.nt.2
GAAAAAAAAAAAGAAAAAAAACAAAAAACAAAAAATACATCACAATAAGAGGCCGAATGTAAATAAAGAC
GAGGGACGCAGAGGGGTGGCCGTGGCTGAGAGGAGACAGCGCCGCAGCACTGAGGGTTTGGGCTTGCAGG
CGCTGCAGGAGACGCCCAGGCGGAGTCTTGTCTCGCAGCCAGCTCTGAGCGGGAGGCCTGAGCGGGAAGC
ATTGGCGTCCGAGCGACTTCTAGGAGCCTGGGGTTCGGCGCTATGGAGGAGCTCGATGGCGAGCCAACAG
TCACTGTAAGGGTACCCCGAACAGGCTTGCTCGTCCTTGCGGGTTGAGAACTGCGTCTGCTTAGTTACTT
CAGGCTTGTCTGCTTCCCTAGTGGTCGCGAGGCGCTCGTCCCCTTCCCTCGACTCAGTTGCCACTTTTCC
GGAGGTCGCAGTGTTAACGAGTTGATTCCAGGCGTGAATTCCAAGAAGAACCAAATGTATTTTGACTGGG
GTCCAGGGGAGATGCTGGTATGTGAAACCTCCTTCAACAAAAAAGAAAAATCAGAGATGGTGCCAAGTTG
CCCCTTTATCTATATCATCCGTAAGGATGTAGATGTTTACTCTCAAATCTTGAGAAAACTCTTCAATGAA
TCCCATGGAATCTTTCTGGGCCTCCAGAGAATTGACGAAGAGTTGACTGGAAAATCCAGAAAATCTCAAT
TGGTTCGAGTGAGTAAAAACTACCGATCAGTCATCAGAGCATGTATGGAGGAAATGCACCAGGTTGCAAT
TGCTGCTAAAGATCCAGCCAATGGCCGCCAGTTCAGCAGCCAGGTCTCCATTTTGTCAGCAATGGAGCTC
ATCTGGAACCTGTGTGAGATTCTTTTTTATTGAAGTGGCCCCAGCTGGCCCTCTCCTCCTCCATCTCCTTG
ACTGGGTCCGGCTCCATGTGTGCGAGGTGGACAGTTTGTCGGCAGATGTTCTGGGCAGTGAGAATCCAAG
CAAACATGACAGCTTCTGGAACTTGGTGACCATCTTGGTGCTGCAGGGCCGGCTGGATGAGGCCCGACAG
ATGCTCTCCAAGGAAGCCGATGCCAGCCCCGCCTCTGCAGGCATATGCCGAATCATGGGGGGACCTGATGA
GGACAATGCCCATTCTTAGTCCTGGGAACACCCAGACACTGACAGACTGGGAGCTGAAGTGGCAGCACTG
GCACGAGGAATGTGAGCGGTACCTCCAGGACAGCACATTCGCCACCAGCCCTCACCTGGAGTCTCTCTTG
AAGATTATGCTGGGAGACGAAGCTGCCTTGTTAGAGCAGAAGGAACTTCTGAGTAATTGGTATCATTTCC
TAGTGACTCGGCTCTTGTACTCCAATCCCACAGTAAAACCCATTGATCTGCACTACTATGCCCAGTCCAG
CCTGGACCTGTTTCTGGGAGGTGAGAGCAGCCCAGAACCCCTGGACAACATCTTGTTGGCAGCCTTTGAG
TTTGACATCCATCAAGTAATCAAAGAGTGCAGCTTCCTCTGAAGACGGGCCAGTTCCTTGCTGTCTGGCA
AGAGGAAACAGCTGGGGTGCACTTCACTGGGAGTTGGGCAAGATGTCGGCAGTTTCCTGGTGCCCTTCAA
GTCCTCCAAAAGTATCGAGCTAAAAGCATCGCCCTGAGCAACTGGTGGTTTGTGGCCCACCTGACAGACC
TGCTGGACCACTGCAAGCTCCTCCAGTCACACAACCTCTATTTCGGTTCCAACATGAGAGAGTTCCTCCT
GCTGGAGTACGCCTCGGGACTGTTTGCTCATCCCAGCCTGTGGCAGCTGGGGGTCGATTACTTTGATTAC
TGCCCCGAGCTGGGCCGAGTCTCCCTGGAGCTGCACATTGAGCGGATACCTCTGAACACCGAGCAGAAAG
CCCTGAAGGTGCTGCGGATCTGTGAGCAGCGGCAGGATGACTGAACAAGTTCGCAGCATTTGTAAGATCTT
AGCCATGAAAGCCGTCCGCAACAATCGCCTGGGTTCTGCCCTCTCTTGGAGCATCCGTGCTAAGGATGCC
GCCTTTGCCACGCTCGTGTCAGACAGGTTCCTCAGGGATTACTGTGAGCGAGGCTGCTTTTCTGATTTGG
ATCTCATTGACAACCTGGGGCCAGCCATGATGCTCAGTGACCGACTGACATTCCTGGGAAAGTATCGCGA
GTTCCACCGTATGTACGGGGAGAAGCGTTTTGCCGACGCAGCTTCTCTCCTTCTGTCCTTGATGACGTCT
CGGATTGCCCCTCGGTCTTTCTGGATGACTCTGCTGACAGATGCCTTGCCCCTTTTGGAACAGAAACAGG
TGATTTTCTCAGCAGAACAGACTTATGAGTTGATGCGGTGTCTGGAGGACTTGACGTCAAGAAGACCTGT
GCATGGAGAATCTGATACCGAGCAGCTCCAGGATGATGACATAGAGACCACCAAGGTGGAAATGCTGAGA
CTTTCTCTGGCACGAAATCTTGCTCGGGCAATTATAAGAGAAGGCTCACTGGAAGGTTCCTGAGAACTGC
TTCAATGTGGTATCTTTGTATGGCAATGTATATAGATTTTTTAAAAGAATAAATGTTGTTTTGCAAATG

>DEX0367_029.nt.3
GAAAAAAAAAAAGAAAAAAAACAAAAAACAAAAAATACATCACAATAAGAGGCCGAATGTAAATAAAGAC
GAGGGACGCAGAGGGGTGGCCGTGGCTGAGAGGAGACAGCGCCGCAGCACTGAGGGTTTGGGCTTGCAGG
CGCTGCAGGAGACGCCCAGGCGGAGTCTTGTCTCGCAGCCAGCTCTGAGCGGGAGGCCTGAGCGGGAAGC
ATTGGCGTCCGAGCGACTTCTAGGAGCCTGGGGTTCGGCGCTATGGAGGAGCTCGATGGCGAGCCAACAG
TCACTGTAAGGGTACCCCGAACAGGCTTGCTCGTCCTTGCGGGTTGAGAACTGCGTCTGCTTAGTTACTT
CAGGCTTGTCTGCTTCCCTAGTGGTCGCGAGGCGCTCGTCCCCTTCCCTCGACTCAGTTGCCACTTTTCC
GGAGGTCGCAGTGTTAACGAGTTGATTCCAGGCGTGAATTCCAAGAAGAACCAAATGTATTTTGACTGGG
GTCCAGGGGAGATGCTGGTATGTGAAACCTCCTTCAACAAAAAAGAAAAATCAGAGATGGTGCCAAGTTG
CCCCTTTATCTATATCATCCGTAAGGATGTAGATGTTTACTCTCAAATCTTGAGAAAACTCTTCAATGAA
TCCCATGGAATCTTTCTGGGCCTCCAGAGAATTGACGAAGAGTTGACTGGAAAATCCAGAAAATCTCAAT
TGGTTCGAGTGAGTAAAAACTACCGATCAGTCATCAGAGCATGTATGGAGGAAATGCACCAGGTTGCAAT
TGCTGCTAAAGATCCAGCCAATGGCCGCCAGTTCAGCAGCCAGGTCTCCATTTTGTCAGCAATGGAGCTC

```
ATCTGGAACCTGTGTGAGATTCTTTTTATTGAAGTGGCCCCAGCTGGCCCTCTCCTCCTCCATCTCCTTG
ACTGGGTCCGGCTCCATGTGTGCGAGGTGGACAGTTTGTCGGCAGATGTTCTGGGCAGTGAGAATCCAAG
CAAACATGACAGCTTCTGGAACTTGGTGACCATCTTGGTGCTGCAGGGCCGGCTGGATGAGGCCCGACAG
ATGCTCTCCAAGGAAGCCGATGCCAGCCCCGCCTCTGCAGGCATATGCCGAATCATGGGGGGACCTGATGA
```
5
```
GGACAATGCCCATTCTTAGTCCTGGGAACACCCAGACACTGACAGAGCTGGAGCTGAAGTGGCAGCACTG
GCACGAGGAATGTGAGCGGTACCTCCAGGACAGCACATTCGCCACCAGCCCTCACCTGGAGTCTCTCTTG
AAGATTATGCTGGGAGACGAAGCTGCCTTGTTAGAGCAGAAGGAACTTCTGAGTAATTGGTATCATTTCC
TAGTGACTCGGCTCTTGTACTCCAATCCCACAGTAAAACCCATTGATCTGCACTACTATGCCCAGTCCAG
CCTGGACCTGTTTCTGGGAGGTGAGAGCAGCCCAGAACCCCTGGACAACATCTTGTTGGCAGCCTTTGAG
```
10
```
TTTGACATCCATCAAGTAATCAAAGAGTGCAGGAATAAAACAGATTTGAGCAGAAGAAGTCTGTTGGATG
CAGGGAGTATAAAAGGGGAAAGCATCCTGCTGTTTCCAGTGGCAGAAGAGAAAGAGAAGTACCATGAAGA
GGGGTGAACAACATCACTCTCGTGCCGGACACAGACAAGGAATGATTAGGCTTGCAGTCCCCCAGTCTCA
GCCTCTCTCCACTCTCACAAGCAGGAAGGGATCTGAGCTCTGTTAGTGGGAAATCTTGGAGTTTGCATTT
CTCTGATCACGAGGCTACCTCCTATCCTGTTTTGAAAGAGGAACAGAAAAGCAGAGCTTGTCCCAGGAAT
```
15
```
TGTGAAATGGGCTGCAATGTCTTCAGAGGCAGTAGCAGTTTTAGGAAGTGTGCCCTCCTGCCTCTGGGGA
AAGGTTGAATTGGGGTCTTGTGGGCACTTGTTTTCTCATGGGATGTATCAGTCAGACCCATGCACAGTTC
TGGGCGGCAGGGAAGTGTGTCTCTTTCTGGACTTGAGCCAGTGAGCTGTTGGCATGCCTTTTTAAAGGCA
TTAGAAAAAAGGTTTCTTTCTTTTTTTTTCCTTTCCTGTTCACACCACCTTCATATGAACATGGAGAAGGG
AGGTTTCTAAGTGGGCCCCTGATTTACTCCTATGTCAATAATCGTGCCAGACCCTAAGGGGCAGAAAAGG
```
20
```
CAGCCAGGGGATCCTGACTCAGAGGCTGTAGTTCGTCAGCCCTGCTAAAGGGCCAGGGGTTTGTCTCTCA
GTGCCTGGGTGGGCATGCTTGACTGCTAACCCCTTAGCTGAAAGGAGATGGCAGGAGTCCTCCACAGGGT
CCCTTCTTCCCTCCTGGCTTACAGCTTCCTCTGAAGACGGGCCAGTTCCTTGCTGTCTGGCAAGAGGGAAA
CAGCTGGGGTGCACTTCACTGGGAGTTGGGCAAGATGTCGGCAGTTTCCTGGTGCCCTTCAAGTCCTCCA
AAAGTATCGAGCTAAAAGCATCGCCCTGAGCAACTGGTGGTTTGTGGCCCACCTGACAGACCTGCTGGAC
```
25
```
CACTGCAAGCTCCTCCAGTCACACAACCTCTATTTCGGTTCCAACATGAGAGAGTTCCTCCTGCTGGAGT
ACGCCTCGGGACTGTTTGCTCATCCCAGRCCTGTGGCTGCCTGGGGGTCCGGCTGTCTGGTACTAGCTTT
TTCATACCGGCGCCTCGCGGGGGCCCTTTTACGCG
```

```
>DEX0367_029.nt.4
```
30
```
GAAAAAAAAAAAGAAAAAAAACAAAAAACAAAAAATACATCACAATAAGAGGCCGAATGTAAATAAAGAC
GAGGGACGCAGAGGGGTGGCCCGTGGCTGAGAGGAGACAGCGCCGCAGCACTGAGGGTTTGGGCTTGCAGG
CGCTGCAGGAGACGCCCAGGCGGAGTCTTGTCTCGCAGCCAGCTCTGAGCGGGAGGCCTGAGCGGGAAGC
ATTGGCGTCCGAGCGACTTCTAGGAGCCTGGGGTTCGGCGCTATGGAGGAGCTCGATGGCGAGCCAACAG
TCACTGTAAGGGTACCCCGAACAGGCTTGCTCGTCCTTGCGGGTTGAGAACTGCGTCTGCTTAGTTACTT
```
35
```
CAGGCTTGTCTGCTTCCCTAGTGGTCGCGAGGCGCTCGTCCCCTTCCCTCGACTCAGTTGCCACTTTTCC
GGAGGTCGCAGTGTTAACGAGTTGATTCCAGGCGTGAATTCCAAGAAGAACCAAATGTATTTTGACTGGG
GTCCAGGGGAGATGCTGGTATGTGAAACCTCCTTCAACAAAAAAGAAAAATCAGAGATGGTGCCAAGTTG
CCCCTTTATCTATATCATCCGTAAGGATGTAGATGTTTACTCTCAAATCTTGAGAAAACTCTTCAATGAA
TCCCATGGAATCTTTCTGGGCCTCCAGAGAATTGACGAAGAGTTGACTGGAAAATCCAGAAAATCTCAAT
```
40
```
TGGTTCGAGTGAGTAAAAACTACCGATCAGTCATCAGAGCATGTATGGAGGAAATGCACCAGGTTGCAAT
TGCTGCTAAAGATCCAGCCAATGGCCGCCAGTTCAGCAGCCAGGTCTCCATTTTGTCAGCAATGGAGCTC
ATCTGGAACCTGTGTGAGATTCTTTTTATTGAAGTGGCCCCAGCTGGCCCTCTCCTCCTCCATCTCCTTG
ACTGGGTCCGGCTCCATGTGTGCGAGGTGGACAGTTTGTCGGCAGATGTTCTGGGCAGTGAGAATCCAAG
CAAACATGACAGCTTCTGGAACTTGGTGACCATCTTGGTGCTGCAGGGCCGGCTGGATGAGGCCCGACAG
```
45
```
ATGCTCTCCAAGGAAGCCGATGCCAGCCCCGCCTCTGCAGGCATATGCCGAATCATGGGGGGACCTGATGA
GGACAATGCCCATTCTTAGTCCTGGGAACACCCAGACACTGACAGAGCTGGAGCTGAAGTGGCAGCACTG
GCACGAGGAATGTGAGCGGTACCTCCAGGACAGCACATTCGCCACCAGCCCTCACCTGGAGTCTCTCTTG
AAGATTATGCTGGGAGACGAAGCTGCCTTGTTAGAGCAGAAGGAACTTCTGAGTAATTGGTATCATTTCC
TAGTGACTCGGCTCTTGTACTCCAATCCCACAGTAAAACCCATTGATCTGCACTACTATGCCCAGTCCAG
```
50
```
CCTGGACCTGTTTCTGGGAGGTGAGAGCAGCCCAGAACCCCTGGACAACATCTTGTTGGCAGCCTTTGAG
TTTGACATCCATCAAGTAATCAAAGAGTGCAGGAATAAAACAGATTTGAGCAGAAGAAGTCTGTTGGATG
CAGGGAGTATAAAAGGGGAAAGCATCCTGCTGTTTCCAGTGGCAGAAGAGAAAGAGAAGTACCATGAAGA
GGGGTGAACAACATCACTCTCGTGCCGGACACAGACAAGGAATGATTAGGCTTGCAGTCCCCCAGTCTCA
GCCTCTCTCCACTCTCACAAGCAGGAAGGGATCTGAGCTCTGTTAGTGGGAAATCTTGGAGTTTGCATTT
```
55
```
CTCTGATCACGAGGCTACCTCCTATCCTGTTTTGAAAGAGGAACAGAAAAGCAGAGCTTGTCCCAGGAAT
TGTGAAATGGGCTGCAATGTCTTCAGAGGCAGTAGCAGTTTTAGGAAGTGTGCCCTCCTGCCTCTGGGGA
AAGGTTGAATTGGGGTCTTGTGGGCACTTGTTTTCTCATGGGATGTATCAGTCAGACCCATGCACAGTTC
TGGGCGGCAGGGAAGTGTGTCTCTTTCTGGACTTGAGCCAGTGAGCTGTTGGCATGCCTTTTTAAAGGCA
TTAGAAAAAAGGTTTCTTTCTTTTTTTTTCCTTTCCTGTTCACACCACCTTCATATGAACGATGATGACAT
```
60
```
AGAGACCACCAAGGTGGAAATGCTGAGACTTTCTCTGGCACGAAATCTTGCTCGGGCAATTATAAGAGAA
```

GGCTCACTGGAAGGTTCCTGAGAACTGCTTCAATGTGGTATCTTTGTATGGCAATGTATATAGATTTTTT
AAAAGAATAAATGTTGTTTTGCAAATG

>DEX0367_030.nt.1
CTCGATGTAACACCCCGACGTCTTCACAAATTAAATTTCAACACATTCTCTATGCTAGTGTGACAAAGCA
GCCCCATAATTTGGTTTTTATTGTTGACCTTTACAGGATGAAGGAGGAGAATCCCCTGTGGCATRCCAAT
GAATCTTTCTGGTGGGGACAGCTGACGAGGCCACGCGGGCAGAGATCCAGATCCGAATCGAGGCCAACGA
GGCCCTGGTGAAGGCCCTGGAGTAGGCGGACATGTGTTATCTACYTCAAGGTAACAAGAGTATGTGGCAA
ACATATACCACCCATAGTGCTTCACAAATGCACTTCTATTTAGCCAGCGTTTATTGTAGTAAACTATTCT
TAATAGAAGTGCATTTTGTGAAGCACACCTCGGCCGCGACCACATCCATCACACTGCGCCGCTCG

>DEX0367_031.nt.1
ACGTGCATGTTGAATCAATGCAAGAGCGAAATGGTTGCTGTGAAATCCCTTTAAGAAACTACATTATTCA
AGTAATGAAAATCACTCAGTTCTACTTCCAGAGAAGAACTGATGAGAAAGGTCTTTAACAGTAATTGAAT
GGAACAAGAAGATAGTATTTTTCTGGGACTGTGTGTCTCCTTTGGTCTTCTCCCCTCCTTCCTGTCACTG
TTGCATACGGTGGTTTCTGCAATGGCTGATCCAGTTTGCCACATTCCACTCCGGCTGCTGTTTCTCTGAG
ATTTTCTTGAGTTTTTATCAAACCAATCCCCAGATAAATCACAATTGCTCTAGAACACCAGCCCTCAGGG
TAACATGTTAGAAATTAGCTTGAAACCCTCAATTTTGAGGGTTTCCCATGACTTTTTAAACGTTGTCCTT
CAGTGTAGCCCCTTGCTCCTAAGGAAAGATTTATTTTCAAAATAGTTTTCAAATGGATGCTCTTCTGTGA
TGGAGCTAAAAAGGAATTATTTCCATATTAATTAATGATGAATTGAGGATGCCTAAGTATGAAATGGAAG
ACGGTTGTATATTTCTCTCAAATGTCTCTGAAAATGAAGTGATGATTATCATTAACAACAATGTGATAAG
AAATTAATATAATTTAAAGCAGCAAGAACAAAATCCTTCTTTCCGATAGCTGGGGATTTGGAAACTTGGT
CCTGCTCAATGCTGCTGCTACTTCTTGTGTCACTATGAACCCATCTTTTGGCCTCTCTGAGTCTCAGGTA
CACCCCCACGATTAGACTACAAGGCCTTATACCGCTCTGACATTACCAGTCAATAACTGTACAACAGAGT
CTAAGGGAATCCTTTTATCTCTTTGCAGAAAACATGATAATATTATTTTAATTATTTGTAAATGCAGGCT
TATTATGCATGTGTTTAAACAGAAAATGTAGCGTTTCTATTAAATCTTTGGATTAAATATGGTTGTGT
AATGAAGTAAATGTCATGTGTATACACTTAACATTCTAAAAATTTGTTTCTTTTACTTTAGAAAATTCAT
AGAAATGGGTTTCATTGATGAAAAAAGAATAGCCATATGGGGCTGGTCCTATGGAGGATACGTTTCATCA
CTGGCCCTTGCATCTGGAACTGGTCTTTTCAAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAAT
ATTACGCGTCTGTCTACACAGAGAGATTCATGGGTCTCCCAACAAAGGATGATAATCTTGAGCACTATAA
GAATTCAACTGTGATGGCAAGAGCAGAATATTTCAGAAATGTAGACTATCTTCTCATCCACGGAACAGCA
GATGATAATGTGCACTTTCAAAACTCAGCACAGATTGCTAAAGCTCTGGTTAATGCACAAGTGGATTTCC
AGGCAATGTGGTACTCTGACCAGAACCACGGCTTATCCGGCCTGTCCACGAACCACTTATACACCCACAT
GACCCACTTCCTAAAGCAGTGTTTCTCTTTGTCAGACTAAAAACGATGCAGATGCAAGCCTGTATCAGAA
TCTGAAAACCTTATATAAACCCCTCAGACAGTTTGCTTATTTTATTTTTTATGTTGTAAAATGCTAGTAT
AAACAAACAAATTAATGTTGTTCTAAAGGCTGTTAAAAAAAAGATGAGGACTCAGAAGTTCAAGCTAAAT
ATTGTTTACATTTTCTGGTACTCTGTGAAAGAAGAAGAAAAGGGAGTCATGCATTTTGCTTTGGACACAGT
GTTTTATCACCTGTTCATTTGAAGAAAAATAATAAAGTCAGAAGTTCAAGTGCTA

>DEX0367_032.nt.1
CGGGCGGGATCAGCGAGCGTCGCGATGTGATGACGTCAGGCCCCGGACAGGCCGGGAGTGGCGTGCTGGG
CGTGCGCGGCTGCGGTACGGCGTGTTGGTCCCAGCGGTTCAGCTGAGGTAGGGACGTGCTGTAGGCCGGA
ATGTTACCGGCTGTTGGATCTGCGGATGAGGAGGAGGATCCTGCGGAGGAGGATTGTCCTGAATTGGTTC
CCATTGAGACGACGCAAAGCGAGGAGGAGGAAAAGTCTGGCCTCGGCGCCAAGATCCCAGTCACAATTAT
CACCGGGTATTTAGGTGCTGGGAAGACAACACTTCTGAACTATATTTTGACAGAGCAACATAGTAAAAGA
GTAGCGGTCATTTTAAATGAATTTGGGGAAGGAAGTGCGCTGGAGAAATCCTTAGCTGTCAGCCAAGGTG
GAGAGCTCTATGAAGAGTGGCTGGAACTTAGAAACCGGTTGCCTCTGCTGTTCAGTGAAGGACAGTGGCCT
TAGAGCTATTGAGAATTTGATGCAAAAGAAGGGGAAATTTGATTACATACTGTTAGAGACCACTGGATTA
GCAGACCCTGGTGCAGTGGCTTCTATGTTTTGGGTTGATGCTGAATTAGGGAGTGATATTTATCTTGATG
GTATCAACTATTGTGGATTCAAAATATGGATTAAAACATTTAACAGAAGAGAAACCTGATGGCCTTAT
CAATGAAGCTACTAGACAAGTTGCTTTGGCAGAAATTGGGCCCCCCCCCAATTTTTTTTTNCCCTTTTTCC
CNAAAATTTTTTTTTTTAAAAAAAAAATTTTTTTNAAA

>DEX0367_033.nt.1
AAAAAGAAAAAATTGTCACTGGTTATCCTTCAGGTATAAAAAATTACAATTAAAATTGTAATGTTATAACC
TTCCATTCTTTCCCCTATGAATTAAAACTTTTTTTTTTTTAACGAGACTAAAAAAATCAAGTTGATAAAATCT
ATGCCCTGTGGGAACTGGTTAGGTCAGGCTACCCACTCCAGATTTCCAGCAGGCTGGCCTCAGCTCTACC
TACAGCCAGTAGCTTATTAAGGCTCTAGGCGCCAGTGACACCTTCAGCAGGTGGTTAGTGCGCCCTAAG
GACTCTCCGCAGCGTCGCTCAGGTTCACAGAACACGCCCAGGGGCGTGTCCAGCTGTCGTCGGGGAGAGC
CCACCTCCCCGGGGGGTGTGGCTAAGGGACGAGGCAGTTCTCGTCCAGAGCCCAGGTAATCCGGGCGGGA

```
          TCAGCTAGCGTCGCGATGTGATGACGTCAGGCCCCGGCCAGGCCGGGAGTGGCGTGCTGGGCGTGCGCGG
          CTGCGGTACGGCGTGTTGGTCCCAGCGGTTCAGCTGAGGTAGGGACGTGCTGTAGGCCGGAATGTTACCG
          GCTGTTGGATCTGTGGATGAGGAAGAGGATCCTGCGGAGGAGGATTGTCCTGAATTGGTTCCCATTGAGA
          CGACGCAAAGCGAGGAGGAGGAAAAGTCTGGCCTCGGCGCCAAGATCCCAGTCACAATTATCACCGGGTA
    5     TTTAGGTGCTGGGAAGACAACACTTCTGAACTATATTTTGACAGAGCAACATAGTAAAAGAGTAGCGGTC
          ATTTTAAATGAATCTGGGGAAGGAAGTGCGCTGGAGAAATCCTTAGCTGTCAGCCAAGGTGGAGAGCTCT
          ATGAAGAGTGGCTGGAACTTAGAAACGGTTGCCTCTGCTGTTCAGTGAAGTGAGGAATGTGTTACTGTG
          TACATGGTTTACTAGAAATGTTTATTGATTATATTTCCAGCTTTAATTTTCTTGAGTAATTTAACTGAAT
          TTACACAGTTTGCTTCATTGTATTTTCAAACAAATAGAAAATAAACTTATTAGGAAGCATTTTCTTAAAG
   10     TGTTTCTTGCTGTCTTTTCTATCTGCTCTAATGTTTTGGTCCTTTTATTGAGTTTTTATTGCTTTTGATG
          TCAGGGCTTATTTAATCTCTAGTGCATGAAAGTCTCATATGTAAAAAATGATTATTCTGAATTTAATCTG
          TCATTGGTCATACTTCTAAGTGTTCAACCTTATAAAAAAATAAATGACTATCAAAAAAGAAAAACCTTACA
          TTATGTTCTAGTAGTTAAGTTTCCAAGGACAGTGATCACTAGTCTACCATAGACCCTAGAAGAGTTACCC
          AACACATAGTAGCACTCAAATATTTGTTGAATGAATTATAAAAATGACTACTTGTATTGTTAATTTTGTG
   15     TATTCTAGTGAATTAAATCTCTTCGGCATCATTTACTCCCTTAGGTATTTGACTTTGTGTCAAATGTTTT
          GGCAAGGATAAAATTATAACAGACTTTCTTGAACAACCAAAATGTAATCTATTAAGGATTTTCCTTCACT
          TTTGATAAAATAAGAAAAAAGGAATTTAAAAACCTTGCATCCTAATGTAAAATAGAATTATATGGTGTTTA
          ATATCAGTGTCCCTTTAGCTATTATATTAAACTACTATAGTTAATAAATTTTATCATTATTTTGTATGTT
          GGTTTTTAAAAATTTCATAAAGCTATAAAAAGATACTTGGTCAGATAAAGTTTCCTCTGCTTTTAATTTT
   20     AATAAAGTATTATTATGTATATGATTTCTTTTTACCTATTATATATATGCATCTATTGTTTTCTCACTGG
          TAAATATGGGACAGACATTTTGTTAGAAGGTTAGAAGTGAGTTAAATTTTCACATTCCTAAGGATACTTT
          TGTCTCGGGTTGTTGAATACATTTTAAAGTGTTTATAATAATCACTTCAAAATATTTAGGTAATTAACTG
          TAAATTATGTTTTGGTATTCTCCAGGGACAATGGCCTTAGAGCTATTGAGAATTTGATGCAAAAGAAGGG
          GAAATTTGATGACATACTGTTAGAGACCACTGGATTAGCAGACCCTGGTATCATAACTATTGTGGATTCA
   25     AAATATGGATTAAAAGTGAAATACCACTAAAGGCAGAAGAGAATGAAATCATCCCACCCACTGACCCGTT
          CCCAATCTGAGCATATGGATAGGGCCCAAAGCAGCAAAGAATTGGACAACTTAAAATGACAGCTCTCTAT
          TCTGAGGACTTGACACCCGGGCTACTGACATTGTTCTTTTCAGTTCTACAAATGTGACAGCATTTAATGC
          AATAGAAGCAATGAATGAAGTTACCCAGAGAAGCATTTAACAGAAGAGAAACCTGATGGCCTTATCAATG
          AAGCTACTAGGCAAGTTGCTTTGGCAGATATCATTCTCATTAATAAAACAGACTTGGTTCCAGAAGAAGA
   30     TGTAAAGAAATTAAGAACGACACTTAGATCCATAAATGGACTAGGACAAATCTTAGAAACACAAAGATCA
          AGGCTCAGCAGCTGATAGACTCAGCAACAGGCAGCCAGGAGCTCTGAGGCTCACAGCTGGCAGTCTAGTT
          CCACTCAGTCTCTACTTGAGAAATTCTTTCTTTGGAAGTACAGCAGAGGCCTTAGGTGAGTGGCTTGTCT
          GCTATGGCAGAGATTAGAGGTGCTGACAGACTGCCATAAGTATTAGGCAGTAACAGCAGCAGCTGCTTAT
          ATGCATGTGAACAGCTGGGGAATTAATTTGGTATGCATTCTCAGGAGCCACTCATCTGCTGGCAGAGGTA
   35     GCAGAAGAATGCCCTTAGTGTAAGTCCTCTACAACCATACACCAAATGTGCTCCCTGCGTTTCAAATTCC
          ATTGTAGAGAGTCTCTGATAATCTCACTTATACCATGAGCCATTCCTCAGTATCTGTCCTCTTCCTGTTA
          GTGTTCTACAATTCCTTTTCTCCTTAATTTTTCTCCGCTTTACAAAATGTCACACAGACAAGTGCATGATA
          CTTAAACAAGCTTTTAAAAATAATGCTCATAAATAGCTTTGGTTCTGTCATAATATTCGTATTTATAAAC
          ATTTTAAGTCAATTCTCTTCTTTTATTTTCATTTCAGAAATATCCATGTCCTGAATAAAAGTTGTGTCTT
   40     GATTAGTTTATTATGTAACAATTTAGTGTGTTTGACATTTCTAACTTTTATTTCTAACATTTGCTTTATT
          ATAGAACAATAAACATGCAGTGATTGATTTTTCTTACTTCAAGTGGATGAGTGAGCAAGTGACTAAAATC
          TTCTGTGAATTCTTCAGTGTATGGTGCTTGCCAATGCATCTGAGAATCTAGGGACTTTCTGAAATAGTAC
          TTCCTTGCTATGAGGACTGAAGTTGGATTAGAATCCTTTTCAATGAAGATCAGATGTCCTGAGTAGAATT
          CTTACTATTGGGTCCTGAATCTTATATTAAATATTCTCTCAAATTCCTTGAGCCACCGCGCCCAGCCTGT
   45     CCTCTGTGATTTTCTGGGCTTATGTTAAAATTATAACTCAATCACCAGTCTTTATAAATTTGCTTTTTTA
          TATTTAAACCAAACCTAATGCTAATTGTGATATGTTTATTTATTCTCACCTGATTTGAATCATTGGATTCA
          ATTAAATGAGTTTAATTATCATTAAAATAATTCTAAGAGAAATAATGCTATTCGGATGGTGGGAATTTTC
          TTTCTACATGCAGCCCCATTCTGAATGAATGAAATCAAATCACGTGACGATCAGGGTCCTAGAGTAACTT
          AATATTTTGTACATTGGTTATTTGACTCCTCATTTTTATATTACATGTTATATCAAGGGAGGGGGTATAA
   50     AAGAAATACAAAAATTGCAGAGGTATCTGGAATGTACCTATTTGTTAATTCTATTTGTCATTTCTTTGGT
          TTCATCTTTTGAGTAATAAGCTGCTTGGAAAAGTTTCTGTTCTTTAGCTGATTTTTTTAGCTATAAAAATG
          TATTTGAAAAGCTCATAAATTTCAGGATTGAAAAGATAATTGAAAGTTTAAAAAAAAACCTAATTCATTGA
          AGTAATAACCAAATAATTTTCAATCTTGATTCAACTGTGATTCAAATCTTACACCATTTGCCCACTTCTA
          TGAATTTTATGTATAAAATTTTTTAAGAGTCAGAGTTTTTTTTTCTTGATTAATTGGATGTATTTCACAG
   55     AATTTCCAACTGCTCACGTTAGTTTTCTTCCTTTTAGAGTTGATCTCTCTAATGTATTAGATCTTCATGC
          CTTTGATAGTCTCTCTGGAATAAGTTTGCAGAAAAAACTTCAGCATGTGCCAGGAACACAACCTCACCTT
          GATCAGAGTATTGTTACAATCACATTTGACGTACCAGGAAATGCAAAGGAAGAACATCTTAATATGTTTA
          TTCAGAATCTTCTGTGGGAAAAGAATGTGAGAAACAAGGACAATCACTGCAGTGGAGGTCATAAGGCTGAA
          GGTACAGTTTACTGTTGCAGACTTTTGGACAAAGTCCTTTTCTTGGCTGTTATAAAAACTATATTTGGTT
   60     TTAAACAGAAACACTGGTTTTAGTACAAATCATCTTTGTCTACTTTCATTTTTCTTTATTATTTTCATGA
```

CTGAAAAGGAGCTTTGGAAATCACTGCATAAGGCTTGATTTATTTGCACAACTTTCTTTAGGGTTGCAGC
TAGAACAAACCTGTGTGCTTTGAAATGTTACCTTCTGCTCTCTGTTCCCAAGTACAGAGAAATAATGTTG
CAAATCTCACTTCTGCTGAACATTATGCTTCCTGATGCATTTAGCAGACACTAAACATTTGTTGTACTCT
AAACAAAGTTACAAAGGACTAGAAGAATTCTTGTTCTGTATTTAGAAACCCACTCACATTACTTGATATT
5 TGGGTATTTAAGTCATGAAAGGTATTTCTTCTAGGAAGCAGTGATTCTAAAGTGTATGCTTAACCAGTCA
GTTGAGTGTCTACTCTTGTGTGTTCACAAGTGTGCACAAAGTTTTTGGTAAATTAAGAATATTATTTCAA
ATAAATTAATTTCATCCCCATAGGAGCCAGTTTATCAGATAATTCGTTTCTCATTTCTGCAAATCAATAC
ACAATGAGCTCATATTCCAGATAAATATAATAGTTTTTCTTTATTTCAACATTGTTCATTGATTGCAACC
CATTTCTAAAACAAATTATTTAATATAACCAAACTCCTTTAACTTCTCAACTTTTCCCACTAAAACTGTG
10 AAACATTATATAAATCTTTAATCAATAGGATATGATACATATTTCATCATTTTGTTTCAGGTTTTGGTT
GTGATAAATAACACTGAAAAACCATCCTAATATATGTATTAACTTTATCATTAGAAAAGGGGATTGTTCA
AAAAAAAATCAATACACAGAAACCTTTTTTGAAAATTAAAAACATTAAGAAACTTCTTGTAATTCACAAA
TAATTGGAATTAAGTAATTTTCTCAGCTTAAGAGTTTGTATTTTTTAAGTGAATATCAGTAATTAATCCA
TTAACATTGATGCTCTATTTCTGCTTTTAAGTTTGAACTGAGGCTGATTAATGTTCAGTCAATTCTCCTT
15 TGAGCAGGATTGGTGTCAATCAAAGACAAATCACAACAAGTGATTGTCCAGGGTGTCCATGAGCTCTATG
ATCTGGAGGAGACTCCAGTGAGCTGGAAGGATGACACTGAGAGAACAAATCGATTGGTCCTCATTGGCAG
AAATTTAGATAAGGATATCCTTAAACAGCTGTTTATAGCTACTGTGACAGAAACAGAAAGCAGTGGACA
ACACATTTCAAAGAAGATCAAGTTTGTACATAACACTAGAGGCATTTCTTATCAAAAGGATTGGATAATA
AAAATAAGTTTCTACTGGGTATATTTCAAGCATTTATTTATTACTTTAGTTACGAATTCCAATATACTTT
20 AAAATGGTATTTGTTTTACAGCATACATAAAATGTAGCAAATCAGTACTGTAAAACATTTAACATTCATA
CAATTATATATAATATCCTTTTTTTTTAAAGAATGGTATTTCACAAAAATATCTTTTGAAATTGGCTTTGG
AGTTTACATATACTGAACATGAAAGTTTATAATAATGATGATACAACTTTCAACATTGTCATTTTTTCTT
AGAACTTCAGCTGATTGCAGAGATATAATGATTACATTGTTATTAAATTTTTTTAACACAAGTAAGTGTC
ACCATTTTATGACATGAAATAAAAGGTTATGACTGTTAAAAAAAAAAATGGCCTCTTGGCCCTACGGTGG
25 GTCTCGGCCCTTCCTTTATGGGGGTATTCATGACCCTCTCCTTGCCCCGCA

>DEX0367_033.nt.2
CGCAATCAGTATTTTTCATACATATTTAAAATTTTCCATCCATACAAAGGTTGCTAACATATTTATTTAAA
ATTTTAAAAATAATTTTATATGCCTCTTGTAACCAGTAATAGGATGCAAAGTTTTTTTTTTAAGGTAACAT
30 TGCTCCTCTTCCCCCATTGTCTCTGTCATCTATAGTACTCCCTGAGGTAGTTCATGTTATTGCAAGGGTG
TGTATCATCACAAACATATTGCTTCCACATACATATATCAGCATACATATATATATATGAGCACATATTA
ATATATGCCCATGCACATACAAGCATGTTTTTGTTAATTTTTACTGAAATGGAATCATACTACATACTTA
TGAACAAATTGCTTCTTTCACCTTATATCACTTATAATTGTTATTCTACAAATCAGTACATATAAAAAAT
ACATATTTACATATGTAATACTCTATCTTTCTGTTATAATTTTCTGCATGCACATGTATGCATATACATA
35 TATTCATAAACATAAAACTTAATCTAAACGGGATTTTTTTTTGGGGGGGTGATATGGTTTGGCTGTGTCC
CCACCCAAATCTCATCTTGAATTCCCCATGTGTTGTTGGGGGGACCCGGTGGGAGGTAATTGAATTATGGGG
GCAGGTCTTTCCTGTGCTGTTCTTGGGATAGTGAAAAAGTCTCACAAGATCTGATGGTTTTAAAAAGGTG
AGTTTCCCTGCATAATCTCTCTTCTCCTGTCTGGTTCCATGTGAGATGTGCCTTTTGCCTTCCGCCATGA
TTGTGAGGCCTCCCCAGCCATGCGGACCCGCCCATAGGTGGAAGGGACTTGCCTTGTCTCGGATGAGACT
40 TTGGACTTTGGACTGTGGACTTTTGAGTTAACACTGAAATGAATTAATGCTGAAATGAATTAAGACTTTG
GGAGACGGTTGAGAAGCCATGACTCATTTTAAAATGTGAGACATGCCTTTCACCTTCTGCCATGATTGTG
AGGCCTCTCCAGCCATGTGGAACTGTGAACACACAGTGTTCATCCCCTCAGGAGGACACAGCTTTTGAGG
TACCATCTTGGAAGCAGAGACCAGCCCTCACTAGACACAAACTTCCTGGCACGTTGATTTTGGACTTCAT
AGCCTCTAGAACTGTGCTGGGAAGACAACACTTCTGAACTATATTTTGACAGAGCAACATAGTAAAAGAG
45 TAGCGGTCATTTTAAATGAATCTGGGGAAGGAAGTGCGCTGGAGAAATCCTTAGCTGTCAGCCAAGGTGG
AGAGCTCTATGAAGAGTGGCTGGAACTTAGAAACGGTTGCCTCTGCTGTTCAGTGAAGTGAGGAATGTGT
TTACTGTGTACATGGTTTACTAGAAATGTTTATTGATTATATTTCCGCTTTAATTTTCTTGAGTAATTT
AACTGAATTTACACAGTTTGTTTCATTGTATTTTCAAACAAATAGAAAATAAACTTATTAGGAAGCATTT
TCTTAAAGTGTTTCTTGCTGTCTTTTCTATCTGCTCTAATGTTTTGGTCCTTTTATTGAGTTTTTATTGC
50 TTTTGATGTCAGGGCTTATTTAATCTCTAGTGCATGAAAGTCTCATATGTAAAAAATGATTATTCTGAAT
TTAATCTGTCATTGGTCATACTTCTAAGTGTTCAACCTTATAAAAAATAAATGACTATCAAAAAAGAAAA
ACCTTACATTATGTTCTAGTAGTTAAGTTTCCAAGGACAGTGATCACTAGTCTACCATAGACCCTAGAAG
AGTTACCCAACACATAGTAGCACTCAAATATTTGTTGAATGAATTATAAAAATGACTACTTGTATTGTTA
ATTTTGTGTATTCTAGTGAATTAAATCTCTTCGGCATCATTTACTCCCTTAGGTATTTGACTTTGTGTCA
55 AATGTTTTGGCAAGGATAAAATTATAACAGACTTTCTTGAACAACCAAAATGTAATCTATTAAGGATTTT
CCTTCACTTTTGATAAAATAAGAAAAAAGGAATTTAAAACCTTGCATCCTAATGTAAAATAGAATTATAT
GGTGTTTAATATCAGTGTCCCTTTAGCTATTATATTAAACTACTATAGTTAATAAATTTTATCATTATTT
TGTATGTTGGTTTTTAAAAATTTCATAAAGCTATAAAAAGATACTTGGTCAGATAAAGTTTCCTCTGCTT
TTAATTTTAATAAAGTATTATTATGTATATGATTTCTTTTTTACCTATTATATATATGCATCTATTGTTTT
60 CTCACTGGTAAATATGGGACAGACATTTTGTTAGAAGGTTAGAAGTGAGTTAAATTTTTCACATTCCTAAG

```
GATACTTTTGTCTCGGGTTGTTGAATACATTTTAAAGTGTTTATAATAATCACTTCAAAATATTTAGGTA
ATTAACTGTAAATTATGTTTTGGTATTCTCCAGGGACAATGGCCTTAGAGCTATTGAGAATTTGATGCAA
AAGAAGGGGAAATTTGATGACATACTGTTAGAGACCACTGGATTAGCAGACCCTGGTATCATAACTATTG
TGGATTCAAAATATGGATTAAAAGTGAAATACCACTAAAGGCAGAAGAGAATGAAATCATCCCACCCACT
GACCCGTTCCCAATCTGAGCATATGGATAGGGGCCCAAAGCAGCAAAGAATTGGACAACTTAAAATGACAG
CTCTCTATTCTGAGGACTTGACACCCGGGCTACTGACATTGTTCTTTTCAGTTCTACAAATGTGACAGCA
TTTAATGCAATAGAAGCAATGAATGAAGTTACCCAGAGAAGCATTTAACAGAAGAGAAACCTGATGGCCT
TATCAATGAAGCTACTAGGCAAGTTGCTTTGGCAGATATCATTCTCATTAATAAAACAGACTTGGTTCCA
GAAGAAGATGTAAAGAAATTAAGAACGACACTTAGATCCATAAATGGACTAGGACAAATCTTAGAAACAC
AAAGATCAAGGCTCAGCAGCTGATAGACTCAGCAACAGGCAGCCAGGAGCTCTGAGGCTCACAGCTGGCA
GTCTAGTTCCACTCAGTCTCTACTTGAGAAATTCTTTCTTTGGAAGTACAGCAGAGGCCTTAGGTGAGTG
GCTTGTCTGCTATGGCAGAGATTAGAGGTGCTGACAGACTGCCATAAGTATTAGGCAGTAACAGCAGCAG
CTGCTTATATGCATGTGAACAGCTGGGGAATTAATTTGGTATGCATTCTCAGGAGCCACTCATCTGCTGG
CAGAGGTAGCAGAAGAATGCCCTTAGTGTAAGTCCTCTACAACCATACACCAAATGTGCTCCCTGCGTTT
CAAATTCCATTGTAGAGAGTCTCTGATAATCTCACTTATACCATGAGCCATTCCTCAGTATCTGTCCTCT
TCCTGTTAGTGTTCTACAATTCCTTTCTCCTTAATTTTTCTCCGCTTTACAAAATGTCACACAGACAAGT
GCATGATACTTAAACAAGCTTTTAAAAATAATGCTCATAAATAGCTTTGGTTCTGTCATAATATTCGTAT
TTATAAACATTTTAAGTCAATTCTCTTCTTTTATTTTCATTTCAGAAATATCCATGTCCTGAATAAAAGT
TGTGTCTTGATTAGTTTATTATGTAACAATTTAGTGTGTTTGACATTTCTAACTTTTATTTCTAACATTT
GCTTTATTATAGAACAATAAACATGCAGTGATTGATTTTTCTTACTTCAAGTGGATGAGTGAGCAAGTGA
CTAAAATCTTCTGTGAATTCTTCAGTGTATGGTGCTTGCCAATGCATCTGAGAATCTAGGGACTTTCTGA
AATAGTACTTCCTTGCTATGAGGACTGAAGTTGGATTAGAATCCTTTTCAATGAAGATCAGATGTCCTGA
GTAGAATTCTTACTATTGGGTCCTGAATCTTATATTAAATATTCTCTCAAATTCCTTGAGCCACCGCGCC
CAGCCTGTCCTCTGTGATTTTCTGGGCTTATGTTAAAATTATAACTCAATCACCAGTCTTTATAAATTTG
CTTTTTTATATTTAAACCAAACCTAATGCTAATTGTGATATGTTATTTATTCTCACCTGATTTGAATCAT
TGGATTCAATTAAATGAGTTTAATTATCATTAAATAATTCTAAGAGAAATAATGTCTATTCGGATGGTGG
GAATTTTCTTTCTACATGCAGCCCCATTCTGAATGAATGAAATCAAATCACGTGACGATCAGGGTCCTAG
AGTAACTTAATATTTTGTACATTGGTTATTTGACTCCTCATTTTTATATTACATGTTATATCAAGGGAGG
GGGTATAAAAGAAATACAAAAATTGCAGAGGTATCTGGAATGTACCTATTTGTTAATTCTATTTGTCATT
TCTTTGGTTTCATCTTTTGAGTAATAAGCTGCTTGGAAAAGTTTCTGTTCTTTAGCTGATTTTTTAGCTA
TAAAAATGTATTTGAAAAGCTCATAAATTTCAGGATTGAAAAGATAATTGAAAGTTTAAAAAAAACCTAA
TTCATTGAAGTAATAACCAAATAATTTTTCAATCTTGATTCAACTGTGATTCAAATCTTACACCATTTGCC
CACTTCTATGAATTTTATGTATAAAATTTTTTAAGAGTCAGAGTTTTTTTTTCTTGATTAATTGGATGTA
TTTCACAGAATTTCCAACTGCTCACGTTAGTTTTCTTCCTTTTAGAGTTGATCTCTCTAATGTATTAGAT
CTTCATGCCTTTGATAGTCTCTCTGGAATAAGTTTGCAGAAAAAACTTCAGCATGTGCCAGGAACACAAC
CTCACCTTGATCAGAGTATTGTTACAATCACATTTGACGTACCAGGAAATGCAAAGGAAGAACATCTTAA
TATGTTTATTCAGAATCTTCTGTGGGAAAAGAATGTGAGAAACAAGGACAATCACTGCATGGAGGTCATA
AGGCTGAAGGGATTGGTGTCAATCAAAGACAAATCACAACAAGTGATTGTCCAGGGTGTCCATGAGCTCT
ATGATCTGGAGGAGACTCCAGTGAGCTGGAAGGATGACACTGAGAGAACAAATCGATTGGTCCTCATTGG
CAGAAATTTAGATAAGGATATCCTTAAACAGCTGTTTATAGCTACTGTGACAGAAACAGAAAAGCAGTGG
ACAACACATTTCAAAGAAGATCAAGTTTGTACATAACACTAGAGGCATTTCTTATCAAAAGGATTGGATA
ATAAAAATAAGTTTCTACTGGGTATATTTCAAGCATTTATTTATTACTTTAGTTACGAATTCCAATATAC
TTTAAAATGGTATTTGTTTTACAGCATACATAAAATGTAGCAAATCAGTACTGTAAAACATTTAACATTC
ATACAATTATATATAATATCCTTTTTTTTTAAAGAATGGTATTTCACAAAAATATCTTTTGAAATTGGCTT
TGGAGTTTACATATACTGAACATGAAAGTTTATAATAATGATGATACAACTTTCAACATTGTCATTTTTT
CTTAGAACTTCAGCTGATTGCAGAGATATAATGATTACATTGTTATTAAATTTTTTTAACACAAGTAAGT
GTCACCATTTTATGACATGAAATAAAAGGTTATGACTGTTAAAAAAAAAAATGGCCTCTTGGCCCTACGG
TGGGTCTCGGCCCTTCCTTTATGGGGGTATTCATGACCCTCTCCTTGCCCCGCA
```

>DEX0367_034.nt.1
```
AGCGGCTTTGCTTCCGCGGCGGGGCCGGAAGTAGGAGCGGCGGCGGCGGCGGCGGCGGTCGAAAGCG
GAGTGAAAGAGGGAGGCAGGGAGCCGGAGAGCCGGAACCGGAGTCGCAGCGGCGGAGACCCCTGTGCGGT
GCGGAGGGGGCGGCGGCCCCGACTCTGACCCGCGCCGGGGGTGGGCCATGGCGGAGATCAGCGACCTGGA
CCGGCAGATCGAGCAGCTGCGTCGCTGCGAGCTCATCAAGGAGAGCGAAGTCAAGGCCCTGTGCGCTAAG
GCCAGAGAGATCTTGGTAGAGGAGAGCAACGTGCAGAGGGTGGACTCGCCAGTCACAGTGTGCGGCGACA
TCCATGGACAATTCTATGACCTCAAAGAGCTGTTCAGAGTAGGTGGCGACGTCCCTGAGACCAACTACCT
CTTCATGGGGACTTTGTGGACCGTGGCTTCTATAGCGTCGAAACGTTCCTCCTGCTGCTGGCACTTAAG
GTTCGCTATCCTGATCGCATCACACTGATCCGGGGCAACCATGAGAGTCGCCAGATCACGCAGGTCTATG
GCTTCTACGATGAGTGCCTGCGCAAGTACGGCTCGGTGACTGTGTGGCGCTACTGCACTGAGATCTTTGA
CTACCTCAGCCTGTCAGCCATCATCGATGGCAAGATCTTCTGCGTGCACGGGGGCCTCTCCCCCTCCATC
```

```
CAGACCCTGGATCAGATTCGGACAATCGACCGAAAGCAAGAGGTGCCTCATGATGGGCCCATGTGTGACC
TCCTCTGGTCTGACCCAGAAGACACCACAGGCTGGGGCGTGAGCCCCCGAGGAGCCGGCTACCTATTTGG
CAGTGACGTGGTGGCCCAGTTCAACGCAGCCAATGACATTGACATGATCTGCCGTGCCCACCAACTGGTG
ATGGAAGGTTACAAGTGGCACTTCAATGAGACGGTGCTCACTGTGTGGTCGGCACCCAACTACTGCTACC
GCTGTGGGAATGTGGCAGCCATCTTGGAGCTGGACGAGCATCTCCAGAAAGATTTCATCATCTTTGAGGC
TGCTCCCCAAGAGACACGGGGCAACCCCGCCCGGCCCCTGCCCCCTCCAACCCTTCTGGCCCTCGCACCA
CTGTGACTCTGCCATCTTCCTCAGACGGAGGCTGGGCGTGGGGGGGGGCTGTCCTGGCTCTGCTGTCCCC
CAAGAGGGTGCTTCGAGGGTGAGGACTTCTCTGGAGAGGCCTGGAGACCTAGCTCCATGTTCCTCCTCCT
CTCTCCCCACTTGAACCATGAAGTTTCCAATAATTTTTTTTTCTTTTTTTCCTTCTTTTTTCTGTTTGTT
TTTAGATAAAAATTTTGAGAAAAAAAATGAAAAAATTCTAATAAAAGAAGAAAAATGGTTTTTGGGTTTG
TGCCAGACTTGGTTGGGAATGCACAGGGTTGCTGGCCTGGGACCTGGGGTGGACGTCAGGGCCTTGACTG
CCTCAGGGCCCCGACCTGCCAAGAGGGCCAGGCTCAGGCCAGTCAGCTTGCCCAGGGCCTCCAGGGGCCG
CTGCCTGCTGGCTTCCCCAGCCCTAGGGAGAGGGCTCTGTTCTCATGGGCAGCCTTATCTTATTAACAGC
ATACACACACAGGCACACACAGAATCACAGGATTCTGATTTCTCCATTTTCCTGAAGAGGAGATGGGC
```

>DEX0367_034.nt.2
```
CATCACAGCCACACACACGCACACGAGGCCTCTGCTCCATCACAGCCACACACATACAGACACACACATG
CAGACACACACATGCAGCCGCTGCTCCACAGTCACTCCGGACCTCCCTGTGACAACACAGGGTGCCCTTC
TGTCTCTGGGAGTGCCCCTTTCACTCTTATGTGCTGTTTGGACGGGGATTGTATGATCACTGTGATTATG
AGGGGGCCCCAGGCATTGGGCGTTCTTTCTCCTCTGGAGCTTATCTGCCCCTACATTTGCTGTCCTGGTC
CCCGCCATCCCACAGACCATATTCAGGCCCCATGCTCTGGTCCCACGGGCCTCTGGTGGACTTGGGGAGG
GGCCAGGCTGCTCACCCTCAAGGGGCAGCGCTGGCAGCCAGAGAAGCCTGAGGACATCCCTCTCCATCCC
TCCCTTTCCGCATCAGACACCACAGGCTGGGGCGTGAGCCCCCGAGGAGCCGGCTACCTATTTGGCAGTG
ACGTGGTGGCCCAGTTCAACGCAGCCAATGACATTGACATGATCTGCCGTGCCCACCAACTGGTGATGGA
AGGTTACAAGTGGCACTTCAATGAGACGGTGCTCACTGTGTGGTCGGCACCCAACTACTGCTACCGGTGA
GCCGGCTGGGCCGGGCTGGGATGGGCGGGCATCTGAGCCGAGCTGCTCCTGACCCTGCTGCCCCGCCTTC
CAGCTGTGGGAATGTGGCAGCCATCTTGGAGCTGGACGAGCATCTCCAGAAAGATTTCATCATCTTTGAG
GCTGCTCCCCAAGAGACACGGGGCATCCCCTCCAACCCTTCTGGCCCTCGCACCACTGTGACTCTGCCATCTTCCTCAGACGGAGG
CTGGGCGTGGGGGGGGGCTGTCCTGGCTCTGCTGTCCCCCAAGAGGGTGCTTCGAGGGTGAGGACTTCTC
TGGAGAGGCCTGGAGACCTAGCTCCATGTTCCTCCTCCTCTCTCCCCACTTGAACCATGAAGTTTCCAAT
AATTTTTTTTTCTTTTTTTTCCTTCTTTTTTTCTGTTTGTTTTTAGATAAAAATTTTGAGAAAAAAAATGAA
AAAATTCTAATAAAAGAAGAAAAATGGTTTTTGGGTTTGTGCCAGACTTGGTTGGGAATGCACAGGGTTG
CTGGCCTGGGACCTGGGGTGGACGTCAGGGCCTTGACTGCCTCAGGGCCCCGACCTGCCAAGAGGGCCAG
GCTCAGGCCAGTCAGCTTGCCCAGGGCCTCCAGGGGCCGCTGCCTGCTGGCTTCCCCAGCCCTAGGGAGA
GGGCTCTGTTCTCATGGGCAGCCTTATCTTATTAACAGCATACACACACAGGCACACACAGAATCACAGG
ATTCTGATTTCTCCATTTTCCTGAAGAGGAGATGGGC
```

>DEX0367_035.nt.1
```
CCGGCCTGGCTCCCAGCGCTCCCTGTCCCCGCCCCGCGGCCAGCTCGCTCCACTCCCACTTCCTGAGCTC
CGCCGTGGGAGCCCTGGAGGCCCGGCCTGGCCGCTCCCGGCCCTGGGGTGCACATCGGCCCTGAGTCCCG
TCCCAGGCTCTGGGCTCGGGCAGCCGCCGCCACCGCTGCCCAGGACGTCGGGCCTCCTGCCTTCCTCCCA
GGCCCCCACGTTGCTGGCCGCCTGGCCGAGTGGCCGCCATGCTCCTGCCTTGGGCCACCTCTGCCCCCGG
CCTGGCCTGGGGGCCTCTGGTGCTGGGCCTCTTCGGGCTCCTGGCAGCATCGCAGCCCCAGGCGGTGCCT
CCATATGCGTCGGAGAACCAGACCTGCAGGGACCAGGAAAAGGAATACTATGAGCCCCAGCACCGCATCT
GCTGCTCCCGCTGCCCGCCAGGCACCTATGTCTCAGCTAAATGTAGCCGCATCCGGGACACAGTTTGTGC
CACATGTGCCGAGAATTCCTACAACGAGCACTGGAACTACCTGACCATCTGCCAGCTGTGCCGCCCCTGT
GACCCAGTGATGGGCCTCGAGGAGATTGCCCCCTGCACAAGCAAACGGAAGACCCAGTGCCGCTGCCAGC
CGGGAATGTTCTGTGCTGCCTGGGCCCTCGAGTGTACACACTGCGAGCTACTTTCTGACTGCCCGCCTGG
CACTGAAGCCGAGCTCAAAGATGAAGTTGGGAAGGGTAACAACCACTGCGTCCCCTGCAAGGCAGGGCAC
TTCCAGAATACCTCCTCCCCCAGCGCCCGCTGCCAGCCCCACACCAGGTGTGAGAACCAAGGTCTGGTGG
AGGCAGCTCCAGGCACTGCCCAGTCCGACACAACCTGCAAAAATCCATTAGAGCCACTGCCCCCAGAGAT
GTCAGAACCTGCCCTGAGCAAGGGGGTGGAGAACTTGCAAGCCCTGCTCTACCAGGCGGCCACAGGCAGT
TCCGAAGCAAGTTTCCCTACTCTCTCACCCCTGTAAACACCCCCACAGGTACAGGTACAGCAGGGCAACC
CTAGCTCCTCTACAGCAGCCCCTCTGTACAATGGCTGAGGCCACAGAAATGTGGCTCTAGCCTGTGTCTG
TTTACAACTCCAAGTCCCACCCTTCCCTACTGTCTCCCAATCCCACTTCCAGACCTTGAAAATCAACTGC
CTAAGCTCCCATCCTGCACACACAAGCCTGCCCAATCCTGGTCTCTGAGCAGGAGGGCACCCACACCACC
CCCAAACATGCCCATCCATGATACAGTCTCTCCTGGCTGCCAAGAGGTCCTCAAGTCCAACTTAGTTCCC
CTTCTCTATAACCCAAGGGAAGTTTCTCTCATTCTGCCTTTAGGAGCTGCATTGTGTCTAGAAGGAAAAA
AGCTGCTCCCTTTTCTCTGTCTGGGTTGCCCAGGGATCTGGAAAGCTCTTCCTTCTCCTCCTCCCCTCTG
```

CCCTTCTTGGGGCTGTGATCACCCTCCTGTCTGCTGTCCTGGCAGGAACCATGCTGATGCTGGCCGTTCT
GCTGCCACTGGCCTTCTTTCTGCTCCTTGCCACCGTCTTCTCCTGCATCTGGAAGAGCCACCCTTCTCTC
TGCAGGAAACTGGGATCGCTGCTCAAGAGGCGTCCGCAGGGAGAGGGACCCAATCCTGTAGCTGGAAGCT
GGGAGCCTCCGAAGGCCCATCCATACTTCCCTGACTTGGTACAGCCACTGCTACCCATTTCTGGAGATGT
TTCCCCAGTATCCACTGGGCTCCCCGCAGCCCCAGTTTTGGAGGCAGGGGTGCCGCAACAGCAGAGTCCT
CTGGACCTGACCAGGGAGCCGCAGTTGGAACCCGGGGAGCAGAGCCAGGTGGCCCACGGTACCAATGGCA
TTCATGTCACCGGCGGGTCTATGACTATCACTGGCAACATCTACATCTACAATGGACCAGTACTGGGGGG
ACCACCGGGTCCTGGAGACCTCCCAGCTACCCCCGAACCTCCATACCCCATTCCCGAAGAGGGGGACCCT
GGCCCTCCCGGGCTCTCTACACCCCACCAGGAAGATGGCAAGGCTTGGCACCTAGCGGAGACAGAGCACT
GTGGTGCCACACCCTCTAACAGGGGCCCAAGGAACCAATTTATCACCCATGACTGACGGAGTCTGAGAAA
AGGCAGAAGAAGGGGGGCACAAGGGCACCTTCTCCCTTGAGGCTGCCCTGCCCACGTGGGATTCACAGGG
GCCTGAGTAGGGCCCGGGGAAGCAGAGCCCTAAGGGATTAAGGCTCAGACACCTCTGAGAGCAGGTGGGC
ACTGGCTGGGTACGGTGCCCTCCACAGGACTCTCCCTACTGCCTGAGCAAACCTGAGGCCTCCCGGCAGA
CCCACCCACCCCCTGGGGCTGCTCAGCCTCAGGCACGGACAGGGCACATGATACCAACTGCTGCCCACTA
CGGCACGCCGCACCGGAGCACGGCACCGAGGGGAGCCGCCACACGGTCACCTGCAAGGACGTCACGGGCCC
CTCTAAAGGATTCGTGGTGCTCATCCCCAAGCTTCAGAGACCCTTTGGGGTTCCACACTTCACGTGGACT
GAGGTAGACCCTGCATGAAGATGAAATTATAGGGAGGACGCTCCTTCCCTCCCCTCCTAGAGGAGAGGAA
AGGGAGTCATTAACAACTAGGGGGTTGGGTAGGATTCCTAGGTATGGGGAAGAGTTTTTGGAAGGGGAGGA
AAATGGCAAGTGTATTTATATTGTAACCACATGCAAATAAAAAGAATGGGACCTA

>DEX0367_035.nt.2
CCGGCCTGGCTCCCAGCGCTCCCTGTCCCCGCCCCGCGGCCAGCTCGCTCCACTCCCACTTCCTGAGCTC
CGCCGTGGGAGCCCTGGAGGCCCGGCCTGGCCGCTCCCGGCCCTGGGGTGCACATCGGCCCTGAGTCCCG
TCCCAGGCTCTGGGCTCGGGCAGCCGCCGCCACCGCTGCCCAGGACGTCGGGCCTCCTGCCTTCCTCCCA
GGCCCCCACGTTGCTGGCCGCCTGGCCGAGTGGCCGCCATGCTCCTGCCTTGGGCCCACCTCTGCCCCCGG
CCTGGCCTGGGGGCCTCTGGTGCTGGGCCTCTTCGGGCTCCTGGCAGCATCGCAGCCCCAGGCGGTGCCT
CCATATGCGTCGGAGAACCAGACCTGCAGGGACCAGGCAGGAAAAGGAATACTATGAGCCCCAGCACCGCATCT
GCTGCTCCCGCTGCCCGCCCAGGCACCTATGTCTCAGCTAAATGTAGCCGCATCCGGGACACAGTTTGTGC
CACATGTGCCGAGAATTCCTACAACGAGCACTGGAACTACCTGACCATCTGCCAGCTGTGCCGCCCCTGT
GACCCAGTGATGGGCCTCGAGGAGATTGCCCCCTGCACAAGCAAACGGAAGACCCAGTGCCGCTGCCAGC
CGGGAATGTTCTGTGCTGCCTGGGCCCTCGAGTGTACACACTGCGAGCTACTTTCTGACTGCCCGCCTGG
CACTGAAGCCGAGCTCAAAGGTCAGAGGTCCCTGAGGGGCTGGATGTGAAAAGGAGGCTGGGTGCCAGGG
ATCTCAAGTGGGAGCAGGGAATATGGTACTGTGTCCACGGCGACCCCCCAGATCCTTGGCTTAAAATTCC
TCTTCTCATTCCATGCAGAGGCTCACAGAAGGGTTCTGGAGGGACATGGAACTGGGACAGGTGCAGGGGA
CACAGGAGGTGGTTCTTCTCAGTACTGGGAAGCTTGGGCAAAATGGTAGGCAGTAGGCCGGGTGTGGTGG
CCCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCAGGAGATCGAGACCA
TCCTGGCTAACACAGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGGGCGTGGTGGTGGGCG
CCTGTATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGTGAACTTGGGAGGCGGAGCTTGCAGTG
AGCCGAGATCGCGCCACTGCACTCCAGCCTGGGCGACAGAGCAAGACTCCGTCTCAAAAAAAAAAAAAAA
AAAGGCAGCAAAGCAGAGCTTGGGAAAGGTGAGAGCTACGCAATGGGGTGGATGGGCATCCTGAGGCTTA
AAGGAAACTCACAGGCCGGCAAAGGGCCCCTCCCTTTTGCCCATTCACCCTGGCTGGCCTGCCTTTCTCT
TGCCAGATGAAGTTGGGAAGGGTAACAACCACTGCGTCCCCTGCAAGGCAGGGCACTTCCAGAATACCTC
CTCCCCCAGCGCCCGCTGCCAGCCCCACACCAGGTGTGAGAACCAAGGTCTGGTGGAGGCAGCTCCAGGC
ACTGCCCAGTCCGACACAACCTGCAAAAATCCATTAGAGCCACTGCCCCCAGAGATGTCAGGAACCATGC
TGATGCTGGCCGTTCTGCTGCCACTGGCCTTCTTTCTGCTCCTTGCCACCGTCTTCTCCTGCATCTGGAA
GAGCCACCCTTCTCTCTGCAGGAAACTGGGATCGCTGCTCAAGAGGCGTCCGCAGGGAGAGGGGACCCAAT
CCTGTAGCTGGAAGCTGGGAGCCTCCGAAGGCCCATCCATACTTCCCTGACTTGGTACAGCCACTGCTAC
CCATTTCTGGAGATGTTTCCCCAGTATCCACTGGGCTCCCCGCAGCCCCAGTTTTTGGAGGCAGGGGTGCC
GCAACAGCAGAGTCCTCTGGACCTGACCAGGGAGCCGCAGTTGGAACCCGGGGAGCAGAGCCAGGTGGCC
CACGGTACCAATGGCATTCATGTCACCGGCGGGTCTATGACTATCACTGGCAACATCTACATCTACAATG
GACCAGTACTGGGGGGACCACCGGGTCCTGGAGACCTCCCAGCTACCCCCGAACCTCCATACCCCATTCC
CGAAGAGGGGGACCCTGGCCCTCCCGGGCTCTCTACACCCCACCAGGAAGATGGCAAGGCTTGGCACCTA
GCGGAGACAGAGCACTGTGGTGCCACACCCTCTAACAGGGGCCCAAGGAACCAATTTATCACCCATGACT
GACGGAGTCTGAGAAAAGGCAGAAGAAGGGGGGCACAAGGGCACCTTCTCCCTTGAGGCTGCCCTGCCCA
CGTGGGATTCACAGGGGCCTGAGTAGGGCCCGGGGAAGCAGAGCCCTAAGGGATTAAGGCTCAGACACCT
CTGAGAGCAGGTGGGCACTGGCTGGGTACGGTGCCCTCCACAGGACTCTCCCTACTGCCTGAGCAAACCT
GAGGCCTCCCGGCAGACCCACCCACCCCCTGGGGCTGCTCAGCCTCAGGCACGGACAGGGCACATGATAC
CAACTGCTGCCCACTACGGCACGCCGCACCGGAGCACGGCACCGAGGGAGCCGCCACACGGTCACCTGCA
AGGACGTCACGGGCCCCTCTAAAGGATTCGTGGTGCTCATCCCCAAGCTTCAGAGACCCTTTGGGGTTCC
ACACTTCACGTGGACTGAGGTAGACCCTGCATGAAGATGAAATTATAGGGAGGACGCTCCTTCCCTCCCC

TCCTAGAGGAGAGGAAAGGGAGTCATTAACAACTAGGGGGTTGGGTAGGATTCCTAGGTATGGGGAAGAG
TTTTGGAAGGGGAGGAAAATGGCAAGTGTATTTATATTGTAACCACATGCAAATAAAAAGAATGGGACCT
A

5 >DEX0367_035.nt.3
AAGAAGAGGGAGGGAGGGGAGAGAGAGAAAAAAGGGTCTGGGGCCCTCATCATTGTTTGGGTCTCCATCTCTT
TCCTGCAGGGAGAGGGACCCAATCCTGTAGCTGGAAGCTGGGAGCCTCCGAAGGCCCATCCATACTTCCC
TGACTTGGTACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCCAGTATCCACTGGGCTCCCCGCAGCC
CCAGTTTTGGAGGCAGGGGTGCCGCAACAGCAGAGTCCTCTGGACCTGACCAGGGAGCCGCAGTTGGAAC
10 CCGGGGAGCAGAGCCAGGTGGCCCACGGTACCAATGGCATTCATGTCACCGGCGGGTCTATGACTATCAC
TGGCAACATCTACATCTACAATGGACCAGTACTGGGGGGGACCACCGGGTCCTGGAGACCTCCCAGCTACC
CCCGAACCTCCATACCCCATTCCCGAAGAGGGGGGACCCTGGCCCTCCCGGGCTCTCTACACCCCACCAGG
AAGATGGCAAGGCTTGGCACCTAGCGGAGACAGAGCACTGTGGTGCCACACCCTCTAACAGGGGCCCAAG
GAACCAATTTATCACCCATGACTGACGGAGTCTGAGAAAAAGGCAGAAGAAGGGGGGCACAAGGGCACCTT
15 CTCCCTTGAGGCTGCCCTGCCCACGTGGGATTCACAGGGGCCTGAGTAGGGCCCGGGGAAGCAGAGCCCT
AAGGGATTAAGGCTCAGACACCTCTGAGAGCCAGGTGGGCACTGGCTGGGTACGGTGCCCTCCACAGGACT
CTCCCTACTGCCTGAGCAATGAGGCCTCCCGGCAGACCCACCCACCCCCTGGGGCTGCTCAGCCTCA
GGCACGGACAGGGCACATGATACCAACTGCTGCCCACTACGGCACGCCGCACCGGAGCACGGCACCGAGG
GAGCCGCCACACGGTCACCTGCAAGGACGTCACGGGCCCCTCTAAAGGATTCGTGGTGCTCATCCCCAAG
20 CTTCAGAGACCCTTTGGGGTTCCACACTTCACGTGGACTGAGGTAGACCCTGCATGAAGATGAAATTATA
GGGAGGACGCTCCTTCCCTCCCCTCCTAGAGGAGAGGAAAGGGAGTCATTAACAACTAGGGGGTTGGGTA
GGATTCCTAGGTATGGGGAAGAGTTTTGGAAGGGGAGGAAAATGGCAAGTGTATTTATATTGTAACCACA
TGCAAATAAAAAGAATGGGACCTA

25 >DEX0367_036.nt.1
TGGGGGGGAGGGGGAAGACGAAGCTTGAAAGACTTGGTAATGGCGACGGGTTTGTCAGAGCACCATAACAT
GGTGTGGGAAGTGAAGACAAATCAGATGCCTAATGCAGTACAGAAACTCCTGTTGGTGATGGACAAGAGA
GCCTCAGGAATGAATGACTCATTGGAGTTGCTGCAGTGTAATGAGAATTTGCCATCTTCACCTGGATATA
ACTCCTGTGTGATGAACACATGGAGCTTGATGACCTTCCTGAACTTCAGGCAGTTCAAAGTGATCCTACCCA
30 ATCTGGCATGTACCAGCTGAGTTCAGATGTTTCACATCAAGAATACCCAAGATCATCTTGGAACCAAAAT
ACCTCAGACATACCAGAAACTACTTACCGTGAAAATGAGGTGGACTGGCTAACAGAATTGGCAAATATCG
CGACCAGTCCACAAAGTCCACTGATGCAGTGCTCATTTTACAATAGATCATCTCCTGTACACATCATAGC
CACTAGCAAAAGTTTACATTCCTATGCACGCCCTCCACCAGTGTCCTCTTCTTCGAAGAGTGAACCAGCC
TTCCCTCATCACCATTGGAAGGAGGAAACACCAGTAAGACACGAAAGGGCAAATAGTGAGTCAGAATCTG
35 GCATTTTCTGCATGTCCTCCCTGTCAGATGATGATGATTTGGGATGGTGCAATTCCTGGCCTTCAACTGT
CTGGCACTGTTTTTTGAAAGCTATGACTTCACACCTATGGATTCTTCTGCAGTTTATGTGTTAAGTAGTA
TGGCTCGCCAGCGTCGTGCATCTTTGTCTTGTGGAGGACCTGGTGGTCAAGACTTTGCAAGATCTGGATT
CAGTAAAAACTGTGGCTCACCTGGATCATCACAGCCTCTTCCCAATTCTTTGTATGCTAAAGCTGTCAAA
AACCACAGCTCAGGGACTGTGAGTGCCACTTCTCCTAATAAGTGCAAAAGACCAATGAATGCCTTCATGC
40 TTTTTGCCAAAAAATACAGAGTTGAATATACTCAGATGTATCCAGGGAAAGATAACAGGTAAAAATAGTG
ATAATTCTGATTACAATAAATGAGTAACACTTCAATACGGTTCATGGTAACTGGAATAGTTAAGTAAACT
TATTCTTTAATAGCTTTTATTCAACTCTTAGGTTGACTTTTTTGTTTGTTTTTGAGACGGAGTCTCGCTG
TGTCACCCAAGCTGGAGTCCAGTGGTACGATCTCAGCTCACTGCAACCTCCGCCTTCCAGGTTCAAAGCC
ATAAGTGTGATCCTTGGTGACAGGTGGAAGAAAATGAAGAATGAAGAGAGAAGAATGTACACATTAGAAG
45 CAAAGGCTTTGGCTGAAGAACAGAAACGTTTAAATCCTGACTGTTGGAAGAGGAAAAGAACCAATTCAGT
ATGTTTCAATAAATAGTGTTTAAAAATTATCTTGCCTTATCCGTGCAGGTGTACTGGCATGTTGGAGCAGT
TGTTACAACCTTTAAGAGGTTGTGTTGATGCTGTCTCTAGCATTTTTATACTTCAAAACCTAATAATGTT
ATCACATCATCATATAAAGACTGCCTTTCCTCAAGGGTCATTTACTCTCATAGCATGGCAAGTCGAAATC
TCAGGCAAACTTCAAAAGAATGTTTGTTCTAAACATTGAGTTGGAAAAATAATTTTGGAATGCAGTGTTT
50 CTAGTTGGGTTGAAAGATAAGCATTATTAATCCATTAATATCACAGTAGAAAATTTCTCTAATAATTTTG
GGATTGTCTCAACAGCTTTTTTTCCTTAATTGGTGAGGGACCAATAAGGCTGGAGGTGTTGGGGTATGTT
AGCTGTCATTCATGGGGACTTGTGCCATTGGCACAGAATTTCTAGATTGAAAAATTTCAAGGAGGAGAAG
AAATAGGCCCAGGAATGGTGGCAAACAAGTCTATATGATAAGCATCTTGGATAGACTTGTTTCTTATAA
CTTTTCCATTACTAAACTAGCATCTTTTAGGTTCAGCGTTTTTAGATTAATTGCAAAAAAATTACATGTG
55 ATATAGTGGGGTCAATTTGAAGTATCATTGAAAGATGAAGGGAAGATATCAATATTTTTATAGTAGGAGG
GCTGGCCAAAGATGTTTTAAATTCATTGTGTACATGTGCCTGTTTCTTAGTAAAATGCAGGTCATATCTG
TTTAAATTTTTCACAATTTCTTTAAAAATAGTGGTCAGTCAGTAGTCAGATGAGAGGGTGGCTGAACTGA
ATTTTGGCAGTTCTTGGTTAATTTAGAATTTCATGTGTTCGGACTTTTCCTGTATCCCTTTTCTGTTCAT
CCTAGTCAATGTCTTATCATTGCATACATTTTCTGGATGCTTGAGCCATCAGATATCAGCTAGCTTTTTA
60 AAATCACCTTTAAGAAAGCAGTATTTTCTAATAGTAAATTTATAAACATCTTATACATCTTAGCTTTCAT

DEX-0367   200   **PATENT**

ATATAGGATTACTATGTATTGATTTGTAAACATTCACTGAGTTTAATTTTATTTCCACAGGGCTCACAAC
AACATTAAACCAGGATGCTTATGTTCTTAAGTCTATATTTGCATATACATTGACTCTTGATGGAAAGACT
TAAGAAGATCAAGGTCTCACCATTTGTCCTCAATTCGTGTGACCATAAGATACTGATAGCATTGAGTCTT
GAAATGATTTAATAATATGAGTGAGGATTTGCTTTCTCCATTAGAGCATTAAGCTAAAACTATCAACATT
5 TTAAACCAAATTGCCTTATTTTTCTTCCAAACTTCATATATGTCTATCAGGTAATAATAGGCTTGAAAAT
TGATATCCTGTGGTGCTAAAGTACAGTAGAAAGAGAGGAGAAGTGTATACATGTTTTATTTTAAATTGTA
CGAAAGGGGAATTTAAAAAAATATGTAACTGCTGTTTATACATTGGCTCCTTACTGCTTATTAATCTGTAT
TGTACACATGATGAAATGAAGCAGAAGCTGGAGTCGGCCTTTCCTCTAGTAACCACCACATGGCTCAGC
ATCTGTGCCAAACATAGGCGCTCCTAGTCTGGTCAGTGCCAAGAGGCTACCAGAACATGGGCAGGTGGC
10 TGGTGTTGGTGTCCCAGCCTAAGAGCCACCTGCTGCAGTTACCATGGCATGCTGAGTTGATGCACCAGGT
GGCAGCAGCCATCCGTTATTATTTCCAATGGAGACCTAGCCCAGGCCAAGGTAAAGTTAGTTAATAGCAT
TGGGATATAGTCACTGTAATGGTGCTATTAACAAACAGTCAACACCATTGTATTTTTTAACTTCGTGTTC
TGTATCTCCTCAGCCATGTATCTTAAATATATTTTGTCATCATAATCTTTATGGTGGGGGCAGACTTTGC
ACTTACTGCAGTGCAACACTTGCACTTTAATTTTCCTCCAACTGTCTAAAATTAGAGCAAATACATTGGC
15 AATACAGCTGCTTTTGCTCTGAGCTACAATCATGGCTTTTCATGTTACTTACCAAGTGGTGTTTCTGGTT
AGGAATCACAGCTGTAAAATTGATTTCAGTTCATCACACTTCTTCATGATGTTGCCCCTAAATTTTGCAC
ACTATATTCTTGTATATTATTTCAAATAAATGGAAAAAAAAGTTGCTTATCTCTGACGTCTACTGATTGA
TTGATTGATTGATTAATCTTGGCTGAATTTTATTTATGAGTTCTCAGAATAACAGGTTCAGGAGATGAGA
TGGAAAACATAAGGCAAATTCCTAACACTGCTAATTAGGAGGCTGCTGATAGGATTTTAAAAAATTAAAAA
20 CAATGTTAATAGTGGAACAAGAGCCATCAGAAAATAAGGACTTAGATTATCTACTATTTATAGAGAATTG
CTATACAATAAAAATTTTTATTTTTCACGTGGCATTTTTGTTACACGGTTAAACAGCCCAAAGGTCTGCT
TTTAATAACATGACAAAGTGTAACAAAAGCCTTGCTTTATTTAAAGATCCAGTATGACATACTTAAAAAG
GCCATTTAACCTTGTATCTTAGCTTGGGATGTGCCAGTGGTTCCTCAACTTTGTTGTCTATTGGAATCAT
TTTGGGGATTTTTTTTTTTTAATAGTGGTGTCTGTCAGGCTCCCACTCCTAGACATGTTAATTAATGTG
25 GAGTGCAAGTTGGGCATCAGGATTTTAAAAAAAATCCCCTGGTGATTGGTATGAATGTGCAGCCAAATTTG
GAAACCACTGAGCTACATGCTTATTTAGTGAAGAAGGCAGGGTTGGGGAAGCAGGGTACATGTTATAAC
TTGATCATATCCCTTTAAAAGAAGTGGGGACTTTGGGTTTATGTACCCATAGGAGGACTTGGTTATGGAT
GGGAAAGACATTTTAGCCTTGTACAAAAATCTGAAAGAGGAAAACATCTTTCTGGAAAAATCAGGTCCAT
GGAATTGAAAGGTGGTGATAACTCAACATTTTTTATGCTAAAGTATAAGTGCTAAAGAGGTAAATAAACG
30 TCGATAGGAAATACCGAACAATCAATTACATTCTTAAAATAGTAGATAGTAGCAGTCTTTTGGAGTATGT
GTTTAACTGCCAACTATGAATGGGTTAGCACAAAGTGGAGATGAACAAAGATTTCATGTCATTACTGAAG
AGCAGACATAGCCTTTTGAAGCAGCTGAACCATTATGGGATAAACTG

>DEX0367_037.nt.1
35 GCGCGGGGAACCGGCGTGGGGTCACGTGGTACCGGCGCATCACGTGGGCGAGTTAGGTGACGCTGCGGGG
CGGGCGGCAGACTGCGGGACGGACGGTGGACGCTGGGACGCGTTTGTAGCTCCGGCCCCGCCGTTCCGA
CCCCCGCCGCCGTCGCCGCCGCATGCGGGGCTAGCACTGCTCTACTCCGGGGTCTTCGTGGCCTTCTGGGC
CTGCGCGCTGGCCGTGGGAGTCTGCTACACCATTTTTTGATTTGGGCTTCCGCTTTGATGTGGCATGGTTC
CTGACGGAGACTTCGCCCTTCATGTGGTCCAACCTGGGCATTGGCCTAGCTATCTCCCTGTCTGTGGTTG
40 GGGCAGCCTGGGGCATCTATATTACCGGCTCCTCCATCATTGGTGGAGGAGTGAAGGCCCCCAGGATCAA
GACCAAGAACCTGGTCAGCATCATCTTCTGTGAGGCTGTGGCCATCTACGGCATCATCATGGCAATTGTC
ATTAGCAACATGGCTGAGCCTTTCAGTGCCACAGACCCCAAGGCCATCGGCCATCGGAACTACCATGCAG
GCTACTCCATGTTTGGGGCTGGCCTCACCGTAGGCCTGTCTAACCTCTTCTGTGGAGTCTGCGTGGGCAT
CGTGGGCAGTGGGGCTGCCCTGGCCGATGCTCAGAACCCCAGCCTCTTTGTAAAGATTCTCATCGTGGAG
45 ATCTTTGGCAGCGCCATTGGCCTCTTTGGGGTCATCGTCGCAATTCTTCAGGTGATGAATCCCCTTGGGA
AGCCTCTGTGTCCTTGTCCCCAACCTAGCCTTACCCTCCTTCTGGAGAAGCTGAAGTGCTCTCCTTCTCT
ACCTATAACTATAGATTCCCCCAAACAGCTTCCACCTCCTCATTTCCATCTTCTGGTTTTCTCCTACAGG
GGCTCTCTATTTTTGTCTCTGATTTGGTGTCACTGACAACAGCCCTGTTCTTTTTCCCCCCCTTTCTAATG
CGTTGTATCTGTATAGCACTTAGTCCTCCCTCTTTCTGTTCTCTCATTTTATCTGCCATGCTTTTCATGT
50 CTCCAGTCTGTATTTGTGATCTAGACTAGAATGGAATTTTGTCTGTCTATATCAGTCTGTCATGGTGTGT
TTGACTGTCTCTCGGAGCTATGTGTATCTCCCTCTCACATGCCTGTCAGTAGGACACCATTGTCTAAATT
GTGTGTGTCTAGTCTGTCTGACAGTGTGTGAGTTTGTGACTAAGGATCAGCCTGAGGGTCTGAGTGTGAC
TGTGGGTGCTTGACTGAGTTGATTCTTAGTGACTAGCTTTATCCATATGCCACTTCTCTCCGTCTGTCCC
ACAGCGTAACTCTGTGGTTGTCTGTCCCTGTCTGCCTTGTCTGTCACAGTGTGTCCATCCTGGTTAGTCA
55 TATATCTCTTGTGGTCTGTCCATCCAACCATGTGCTAGATGTCATCTTCTATCATTTGTCACTGCCTTAC
CTATTCCTCTGTCCCTGCCTGATTTCTCCTGTTCTTTTTTTGCAACCTCCAGAGTGAAGATGGGTGACTA
GATGATATGTGTGGGTGGGGCCGTGCCTCACTTTTATTTATTGCTGGTTTTCCTGGGACAGCTGGAGCTG
TGTCCCTTAGCCTTTCAGAGGCTTGGTGTTCAGGGCCCTCCCTGCACTCCCCTCTTGCTGCGTGTTGATT
TGGAGGCACTGCAGTCCAGGCCGAGTCCTCAGTGCGGGGAGCAGGCTGCTGCTGACTCTGTGCAGCT
60 GCGCACCTGTGTCCCCCACCTCCACCCTCAACCCATCTTCCTAGTGTTTGTGAAATAAACTTGGTATTTG

TCTGGGTCA

>DEX0367_038.nt.1
CAGAAACCAGGGGGGAGAAGGCGGCCGAGCCCCAGCTCTCCGAGCACCGGGTCGGAAGCCGCGACCCGAGC
CGCGCAGGAAGCTGGGACCGGAACCTCGGCGGACCCGGCCCCACCCAACTCACCTGCGCAGGTCACCAGC
ACCCTCGGAACCCAGAGGCCCGCGCTCTGAAGGTGACCCCCCTGGGGAGGAAGGCGATGGCCCCTGCGAG
GACGATGGCCCGCGCCCGCCTCGCCCCGGCCGGCATCCCTGCCGTCGCCTTGTGGCTTCTGTGCACGCTC
GGCCTCCAGGGCACCCAGGCCGGGCCACCGCCCGCGCCCCCTGGGCTGCCCGCGGGAGCCGACTGCCTGA
ACAGCTTTACCGCCGGGGTGCCTGGCTTCGTGCTGGACACCAACGCCTCGGTCAGCAACGGAGCTACCTT
CCTGGAGTCCCCCACCGTGCGCCGGGGCTGGGACTGCGTGCGCGCCTGCTGCACCACCCAGAACTGCAAC
TTGGCGCTAGTGGAGCTGCAGCCCGACCGCGGGGAGGACGCCATCGCCGCCTGCTTCCTCATCAACTGCC
TCTACGAGCAGAACTTCGTGTGCAAGTTCGCGCCCAGGGAGGGCTTCATCAACTACCTCACGAGGGGAAGT
GTACCGCTCCTACCGCCAGCTGCGGCCCAGGGCTTTGGAGGGTCTGGGATCCCCAAGGCCTGGGCAGGC
ATAGACTTGAAGGTACAACCCCAGGAACCCCTGGTGCTGAAGGATGTGGAAAACACAGATTGGCGCCTAC
TGCGGGGTGACACGGATGTCAGGGTAGAGAGGGAAAGACCCAAACCAGGTGGAACTGTGGGGACTCAAGGA
AGGCACCTACCTGTTCCAGCTGACAGTGACTAGCTCAGACCACACCAGAGGACACGGCCAACGTCACAGTC
ACTGTGCTGTCCACCAAGCAGACAGAAGACTACTGCCTCGCATCCAACAAGGTGGGTCGCTGCCGGGGCT
CTTTCCCACGCTGGTACTATGACCCCACGGAGCAGATCTGCAAGAGTTTCGTTTATGGAGGCTGCTTGGG
CAACAAGAACAACTACCTTCGGGAAGAAGAGTGCATTCTAGCCTGTCGGGGTGTGCAAGGTGGGCCTTTG
AGAGGCAGCTCTGGGGCTCAGGCGACTTTCCCCAGGGCCCCTCCATGGAAAGGCGCCATCCAGTGTGCT
CTGGCACCTGTCAGCCCACCCAGTTCCGCTGCAGCAATGGCTGCTGCATCGACAGTTTCCTGGAGTGTGA
CGACACCCCCAACTGCCCCGACGCCTCCGACGAGGCTGCCTGTGAAAAATACACGAGTGGCTTTGACGAG
CTCCAGCGCATCCATTTCCCCAGTGACAAAGGGCACTGCGTGGACCTGCCAGACACAGGACTCTGCAAGG
AGAGCATCCCGCGCTGGTACTACAACCCCTTCAGCGAACACTGCGCCCCGCTTTACCTATGGTGGTTGTTA
TGGCAACAAGAACAACTTTGAGGAAGAGCAGCAGCAGTGCCTCGAGTCTTGTCGCGGCATCTCCAGTGAGTGG
GCCAGTGAGAGGGGTGGGCATGTATGGGGGAAGGCTTAGCCAGTGGCCTCCACTCTGTCCCCAGGCCTTTG
GGAGTCACCCCTCCATCCTCAGAGCGCCTGGAGTAGGAGTGGGAGAGGATGCCAGTGTGAGGTCTGGGGC
ACTGGGGAGCAGCTGAGTCTCTGAGCCCTCAGTACTGACTGGTCTTTCCCCTCATCATTCTGCAGAGAAG
GATGTGTTTGGCCTGAGGCGGGAAATCCCCATTCCCAGCACAGGCTCTGTGGAGATGGCTGTCGCAGTGT
TCCTGGTCATCTGCATTGTGGTGGTGGTAGCCATCTTGGGTTACTGCTTCTTCAAGAACCAGAGAGAAAGGA
CTTCCACGGACACCACCACCACCCACCACCCACCCCTGCCAGCTCCACTGTCTCCACTACCGAGGACACG
GAGCACCTGGTCTATAACCACACCACCCGGCCCCTCTGAGCCTGGGTCTCACCGGCTCTCACCTGGCCCT
GCTTCCTGCTTGCCAAGGCAGAGGCCTGGGCTGGGAAAAACTTTGGAACCAGACTCTTGCCTGTTTCCCA
GGCCCACTGTGCCTCAGAGACCAGGGCTCCAGCCCCTCTTGGAGAAGTCTCAGCTAAGCTCACGTCCTGA
GAAAGCTCAAAGGTTTGGAAGGAGCAGAAAACCCTTGGGCCAGAAGTACCAGACTAGATGGACCTGCCTG
CATAGGAGTTTGGAGGAAGTTGGAGTTTTGTTTCCTCTGTTCAAAGCTGCCTGTCCCTACCCCATGGTGC
TAGGAAGAGGGAGTGGGGTGGTGTCAGACCCTGGAGGCCCCAACCCTGTCCTCCCCGAGCTCCTCTTCCATG
CTGTGCGCCCAGGGCTGGGAGGAAGGACTTCCCTGTGTAGTTTGTGCTGTAAAGAGTTGCTTTTTGTTTA
TTTAATGCTGTGTGGCATGGGTGAAGAGGAGGGGAAGAGGCCTGTTTGGCCTCTCTGTCCTCTCTTCCTCTT
CCCCCAAGATTGAGCTCTCTGCCCTTGATCAGCCCCACCCTGGCCTAGACCAGCAGACAGAGCCAGGAGA
GGCTCAGCTGCATTCCGCAGCCCCCACCCCCAAGGTTCTCCAACATCACAGCCCAGCCCACCCACTGGGT
AATAAAAGTGGTTTGTGGA

>DEX0367_039.nt.1
GGTGGGAGGAAACCGCCTGGAGCCGCCGGGAGTGGACGCCGCCGAGGCCCGGAGTCGCGCCTGCAGGTGC
AGCCCATCAGGGACCCACAGCGCCTGGGAGGATGGTGCGGATCTTGGCCAATGGGGAAATCGTGCAGGAT
GACGACCCCCGAGTGAGGACCACTACCCAGCCACCAAGAGGTAGCATTCCTCGACAGAGCTTCTTCAATA
GGGGCCATGGTGCTCCCCCAGGGGGTCCTGGCCCCCCGCCAGCAGCAGGCAGGTGCCAGGCTGGGTGCTGC
TCAGTCCCCCTTCAATGACCTCAACCGGCAGCTGGTGAACATGGGCTTTCCGCAGTGGCATCTCGGCAAC
CATGCTGTGGAGCCGGTGACCTCCATCCTGCTCCTCTTCCTGCTCATGATGCTTGGTGTTCGTGGCCTCC
TCCTGGTTGGCCTTGTCTACCTGGTGTCCCACCTGAGTCAGCGGTGACCTCTGAGGGCTGATAGGGGTGG
GTTTGTTGAGAGGGACTTGCTGGGCCTTGGTGTGAGAGCAGGCATATTTGGAGGGGATCTGGTGGTGCCT
TGAAGGTATGATCAGAGAGGGGACCACAGGTGTGTGTTTCCCCTTTGTGTTAAGCGTGAGGCAGAGGGAG
ACGTTAGTCCAGCATTTCCAAAGTGTGGGTGGGTCCGTTGGTTCCCGAGATACTTTTAGGTGGTATGGGG
CCTGCATTAAGTGGCACAAAATCAGAGCAAGAAAGCGATGCCCTTCCCAATTCTCTCAATCCTTTTATGC
CGAGAAGATCTCAGCTGGATGCCAACATGTTCCGATGCCTGTGGAAGACATGCCGACGTCTCCTCTGCCT
AGGGAGCAGGACTTGGGCTTAGGGCAGGTGGAAAAAAATTCCAGACTTTTTTAGCACTGTTTTTGTTTTAA
TGGTATATTTTTATTGGCTACTTTATTGTTTAGGACAAGTGGTAGTGGCATTCTATTTATTGTGACCTTT
TCAATAAATAGATTTAAGTAAAAA

>DEX0367_039.nt.2
GGTGGGAGGAAAACCGCCTGGAGCCGCCGGGGAGTGGACGCCGCCGAGGCCCGGGAGTCGCGCCTGCAGACAC
AGCATCTACTCAGCGTGGGTCACCTCTGTGAACATCACTGACTGCAAGCCTCCCTCAATTTCTGGTAAAT
TTTTTCCTTCAAGCTTTTTGAGGTCAGGGGGAAAATCCTCCACAGATCTTCACGAGTCTGGTCCCAGATC
ACTAACAGCTAAAGGTTTTGTGGTGGGGAGAACTTCTGTGCCACTTGGCTTCCATTAGCCTGCCCACCCT
GACCACACAATGCAGGTGCAGCCCATCAGGGACCCACAGCGCCTGGGAGGATGGTGCGGATCTTGGCCAA
TGGGGAAATCGTGCAGGATGACGACCCCCGAGTGAGGACCACTACCCAGCCACCAAGAGGTAGCATTCCT
CGACAGAGCTTCTTCAATAGGGGCCATGGTGCTCCCCAGGGGGTCCTGGCCCCCGCCAGCAGCAGGCAG
GTGCCAGGCTGGGTGCTGCTCAGTCCCCCTTCAATGACCTCAACCGGCAGCTGGTGAACATGGGCTTTCC
GCAGTGGCATCTCGGCAACCATGCTGTGGAGCCGGTGACCTCCATCCTGCTCCTCTTCCTGCTCATGATG
CTTGGTGTTCGTGGCCTCCTCCTGGTTGGCCTTGTCTACCTGGTGTCCCACCTGAGTCAGCGGTGACCTC
TGAGGGCTGATAGGGGTGGGTTTGTTGAGAGGGACTTGCTGGGCCTTGGTGTGAGAGCAGGCATATTTGG
AGGGGATCTGGTGGTGCCTTGAAGGTATGATCAGAGAGGGGACCACAGGTGTGTGTTCCCCTTTGTGTT
AAGCGTGAGGCAGAGGGAGACGTTAGTCCAGCATTTCCAAAGTGTGGGTGGGTCCGTTGGTTCCCGAGAT
ACTTTTAGGTGGTATGGGGCCTGCATTAAGTGGCACAAAATCAGAGCAAGAAAGCGATGCCCTTCCCAAT
TCTCTCAATCCTTTTATGCCGAGAAGATCTCAGCTGGATGCCAACATGTTCCGATGCCTGTGGAAGACAT
GCCGACGTCTCCTCTGCCTAGGGAGCAGGACTTGGGCTTAGGGCAGGTGGAAAAAAATTCCAGACTTTTTT
AGCACTGTTTTTGTTTTAATGGTATATTTTTATTGGCTACTTTATTGTTTAGGACAAGTGGTAGTGGCAT
TCTATTTATTGTGACCTTTTCAATAAATAGATTTAAGTAAAAA

>DEX0367_039.nt.3
GGAGTGATGATGAGTATTAAAAATGCTGGCCCTAGAGCCATGCCCTGTAGTCCAACCTGCCTCAACCACT
TCCTCTCTCCTCTGTCCTGGGCCTGATGTGAGGAATATAGACTGGCATGTTCTGTGCTCTGAGAGCTTGT
GACTTGGTTGTTGGGATTCTTATGGGTGGCATGGGGGATTGCTGAGAAAGGAGGGTGTGACCCTGGGTGG
CTGTAGGAAGGGCAGGATGCAGCCTAGAGAGGTTGATTCCTGCAGGTGCCTCAACCCAATGATTCATTC
ATCTCTCTCTGTCTTCACATTCCAGAGCTTCTTCAATAGGGGCCATGGTGCTCCCCCAGGGGGTCCTGGC
CCCCGCCAGCAGCAGGCAGGTGCCAGGCTGGGTGCTGCTCAGTCCCCCTTCAATGACCTCAACCGGCAGC
TGGTGAACATGGGCTTTCCGCAGTGGCATCTCGGCAACCATGCTGTGGAGCCGGTGACCTCCATCCTGCT
CCTCTTCCTGCTCATGATGCTTGGTGTTCGTGGCCTCCTCCTGGTTGGCCTTGTCTACCTGGTGTCCCAC
CTGAGTCAGCGGTGACCTCTGAGGGCTGATAGGGGTGGGTTTGTTGAGAGGGACTTGCTGGGCCTTGGTG
TGAGAGCAGGCATATTTGGAGGGGATCTGGTGGTGCCTTGAAGGTATGATCAGAGAGGGGACCACAGGTG
TGTGTTTCCCCTTTGTGTTAAGCGTGAGGCAGAGGGAGACGTTAGTCCAGCATTTCCAAAGTGTGGGTGG
GTCCGTTGGTTCCCGAGATACTTTTAGGTGGTATGGGGCCTGCATTAAGTGGCACAAAATCAGAGCAAGA
AAGCGATGCCCTTCCCAATTCTCTCAATCCTTTTATGCCGAGAAGATCTCAGCTGGATGCCAACATGTTC
CGATGCCTGTGGAAGACATGCCGACGTCTCCTCTGCCTAGGGAGCAGGACTTGGGCTTAGGGCAGGTGGA
AAAAATTCCAGACTTTTTTAGCACTGTTTTTGTTTTAATGGTATATTTTTATTGGCTACTTTATTGTTTA
GGACAAGTGGTAGTGGCATTCTATTTATTGTGACCTTTTCAATAAATAGATTTAAGTAAAAA

>DEX0367_040.nt.1
GAAAAAGAGGAAAGAAAGAAAAAGGGAGAATGWTRATACTATAGGCACTGGTTCTCTAATGCATGCTCG
AGCGGCGCAGTGTGATGGATCGCCCGGGCAGGTACCCAGCTGCAGAACTTCACCCTGGACAGGAGCAGTG
TCCTTGTGGATGGGTATTCTCCCAACAGAAATGAGCCCTTAACTGGGAATTCTGACCTTCCCTTCTGGGC
TGTCATCCTCATCGGCTTGGCAGGACTCCTGGGACTCATCACATGCCTGATCTGCGGTGTCCTGGTGACC
ACCCGCCGGCGGAAGAAGGAAGGAGAATACAACGTCCAGCAACAGTGCCCAGGCTACTACCAGTCACACC
TAGACCTGGAGGATCTGCAATGACTGGAACTTGCCGGTGCCTGGGGTGCCTTTCCCCCAGCCAGGGTCCA
AAGAAGCTTGGCTGGGGCAGAAATAAACCATATTGGTCGGTAAAAAAACTCTGAGCCT

>DEX0367_040.nt.2
CGCGTTGATCCCATCGGACCTGGACTGGACAGAGAGCGGCTATACTGGGAGCTGAGCCAGCTGACCAACA
GCATCACAGAGCTGGGACCCTACACCCTGGATAGGGACAGTCTCTATGTCAATGGCTTCAACCCTTGGAG
CTCTGTGCCAACCACCAGCACTCCTGGGACCTCCACAGTGCACCTGGCAACCTCTGGGACTCCATCCTCC
CTGCCTGGCCACACAGCCCCTGTCCCTCTCTTGATACCATTCACCCTCAACTTTACCATCACCAACCTGC
ATTATGAAGAAAACATGCAACACCCTGGTTCCAGGAAGTTCAACACCACGGAGAGGGTTCTGCAGGGTCT
GCTCAAGCCCTTGTTCAAGAGCACCAGCGTTGGCCCTCTGTACTCTGGCTGCAGGCTGACCTTGCTCAGA
CCTGAGAAACATGGGGCAGCCACTGGAGTGGACGCCATCTGCACCCTCCGCCTTGATCCCACTGGTCCTG
GACTGGACAGAGAGCGGCTATACTGGGAGCTGAGCCAGCTGACCAACAGCGTTACAGAGCTGGGCCCCTA
CACCCTGGACAGGGACAGTCTCTATGTCAATGGCTTCACCCATCGGAGCTCTGTGCCAACCACCAGTATT
CCTGGGACCTCTGCAGTGCACCTGGAAACCTCTGGGACTCCAGCCTCCCTCCCTGGCCACACAGCCCCTG
GCCCTCTCCTGGTGCCATTCACCCTCAACTTCACTATCACCAACCTGCAGTATGAGGAGGACATGCGTCA
CCCTGGTTCCAGGAAGTTCAACACCACGGAGAGAGTCCTGCAGGGTCTGCTCAAGCCCTTGTTCAAGAGC

```
ACCAGTGTTGGCCCTCTGTACTCTGGCTGCAGACTGACCTTGCTCAGGCCTGAAAAACGTGGGGCAGCCA
CCGGCGTGGACACCATCTGCACTCACCGCCTTGACCCTCTAAACCCTGGACTGGACAGAGAGCAGCTATA
CTGGGAGCTGAGCAAACTGACCCGTGGCATCATCGAGCTGGGCCCTACCTCCTGGACAGAGGCAGTCTC
TATGTCAATGGTTTCACCCATCGGAACTTTGTGCCCATCACCAGCACTCCTGGGACCTCCACAGTACACC
TAGGAACCTCTGAAACTCCATCCTCCCTACCTAGACCCATAGTGCCTGGCCCTCTCCTGGTGCCATTCAC
CCTCAACTTCACCATCACCAACTTGCAGTATGAGGAGGCCATGCGACACCCTGGCTCCAGGAAGTTCAAT
ACCACGGAGAGGGTCCTACAGGGTCTGCTCAGGCCCTTGTTCAAGAATACCAGTATCGGCCCTCTGTACT
CCAGCTGCAGACTGACCTTGCTCAGGCCAGAGAAGGACAAGGCAGCCACCAGAGTGGATGCCATCTGTAC
CCACCACCCTGACCCTCAAAGCCCTGGACTGAACAGAGAGCAGCTGTACTGGGAGCTGAGCCAGCTGACC
CACGGCATCACTGAGCTGGGCCCCTACACCCTGGACAGGGACAGTCTCTATGTCGATGGTTTCACTCATT
GGAGCCCCATACCAACCACCAGCACTCCTGGGACCTCCATAGTGAACCTGGGAACCTCTGGGATCCCACC
TTCCCTCCCTGAAACTACAGCCACCGGCCCTCTCCTGGTGCCATTCACACTCAACTTCACCATCACTAAC
CTACAGTATGAGGAGAACATGGGTCACCCTGGCTCCAGGAAGTTCAACATCACGGAGAGTGTTCTGCAGG
GTCTGCTCAAGCCCTTGTTCAAGAGCACCAGTGTTGGCCCTCTGTATTCTGGCTGCAGACTGACCTTGCT
CAGGCCTGAGAAGGACGGAGTAGCCACCAGAGTGGACGCCATCTGCACCCACCGCCCTGACCCCAAAATC
CCTGGGCTAGACAGACAGCAGCTATACTGGGAGCTGAGCCAGCTGACCCACAGCATCACTGAGCTGGGAC
CCTACACCCTGGATAGGGGACAGTCTCTATGTCAATGGTTTCACCCAGCGGAGCTCTGTGCCCACCACCAG
CACTCCTGGGACTTTCACAGTACAGCCGGAAACCTCTGAGACTCCATCATCCCTCCCTGGCCCCACAGCC
ACTGGCCCTGTCCTGCTGCCATTCACCCTCAATTTTACCATCATTAACCTGCAGTATGAGGAGGACATGC
ATCGCCCTGGCTCCAGGAAGTTCAACACCACGGAGAGGGTCCTTCAGGGTCTGCTTATGCCCTTGTTCAA
GAACACCAGTGTCAGCTCTCTGTACTCTGGTTGCAGACTGACCTTGCTCAGGCCTGAGAAGGATGGGGCA
GCCACCAGAGTGGATGCTGTCTGCACCCATCGTCCTGACCCCAAAAGCCCTGGACTGGACAGAGAGCGGC
TGTACTGGAAGCTGAGCCAGCTGACCCACGGCATCACTGAGCTGGGCCCCTACACCCTGGACAGGCACAG
TCTCTATGTCAATGGTTTCACCCATCAGAGCTCTATGACGACCACCAGAACTCCTGATACCTCCACAATG
CACCTGGCAACCTCGAGAACTCCAGCCTCCCTGTCTGGACCTACGACCGCCAGCCCTCTCCTGGTGCTAT
TCACAATTAACTTCACCATCACTAACCTGCGGTATGAGGAGAACATGCATCACCCTGTGTTCAAGAACACCAGTGTTGGCCCTCTG
TACTCTGGCTGCAGACTGACCTTGCTCAGGCCCAAGAAGGATGGGGCAGCCACCAAAGTGGATGCCATCT
GCACCTACCGCCCTGATCCCAAAAGCCCTGGACTGGACAGAGAGCAGCTATACTGGGAGCTGAGCCAGCT
AACCCACAGCATCACTGAGCTGGGCCCCTACACCCTGGACAGGGACAGTCTCTATGTCAATGGTTTCACA
CAGCGGAGCTCTGTGCCCACCACTAGCATTCCTGGGACCCCCACAGTGGACCTGGGAACATCTGGGACTC
CAGTTTCTAAACCTGGTCCCTCGGCTGCCAGCCCTCTCCTGGTGCTATTCACTCTCAACTTCACCATCAC
CAACCTGCGGTATGAGGAGAACATGCAGCACCCTGGCTCCAGGAAGTTCAACACCACGGAGAGGGTCCTT
CAGGGCCTGCTCAGGTCCCTGTTCAAGAGCACCAGTGTTGGCCCTCTGTACTCTGGCTGCAGACTGACTT
TGCTCAGGCCTGAAAAGGATGGGACAGCCACTGGAGTGGATGCCATCTGCACCCACCACCCTGACCCCAA
AAGCCCTAGGCTGGACAGAGAGCAGCTGTATTGGGAGCTGAGCCAGCTGACCCACAATATCACTGAGCTG
GGCCCCTATGCCCTGGACAACGACAGCCTCTTTGTCAATGGTTTCACTCATGGAGCTCTGTGTCCACCA
CCAGCACTCCTGGGACCCCCACAGTGTATCTGGGAGCATCTAAGACTCCAGCCTCGATATTTGGCCCTTC
AGCTGCCAGCCATCTCCTGATACTATTCACCCTCAACTTCACCATCACTAACCTGCGGTATGAGGAGAAC
ATGTGGCCTGGCTCCAGGAAGTTCAACACTACAGAGAGGGTCCTTCAGGGCCTGCTAAGGCCCTTGTTCA
AGAACACCAGTGTTGGCCCTCTGTACTCTGGCTGCAGGCTGACCTTGCTCAGGCCAGAGAAAGATGGGGA
AGCCACCGGAGTGGATGCCATCTGCACCCACCGCCCTGACCCCACAGGCCCTGGGCTGGACAGAGAGCAG
CTGTATTTGGAGCTGAGCCAGCTGACCCACAGCATCACTGAGCTGGGCCCCTACACACTGGACAGGGACA
GTCTCTATGTCAATGGTTTCACCCATCGGAGCTCTGTACCCACCACCAGCACCGGGGTGGTCAGCGAGGA
GCCATTCACACTGAACTTCACCATCAACAACCTGCGCTACATGGCGGACATGGGCCAACCCGGCTCCCTC
AAGTTCAACATCACAGACAACGTCATGAAGCACCTGCTCAGTCCTTTGTTCCAGAGGAGCAGCCTGGGTG
CACGGTACACAGGCTGCAGGGTCATCGCACTAAGGTCTGTGAAGAACGGTGCTGAGACACGGGTGGACCT
CCTCTGCACCTACCTGCAGCCCCTCAGCGGCCCAGGTCTGCCTATCAAGCAGGTGTTCCATGAGCTGAGC
CAGCAGACCCATGGCATCACCCGGCTGGGCCCCTACTCTTGGACAAAGACAGCTCTACCTTAACGGTT
ACAATGAACCTGGTCCAGATGAGCTCCTACAACTCCCAAGCCAGCCACCACATTCCTGCCTCCTCTGTC
AGAAGCCACAACAGCCATGGGGTACCACCTGAAGACCCTCACACTCAACTTCACCATCTCCAATCTCCAG
TATTCACCAGATATGGGCAAGGGCTCAGCTACATTCAACTCCACCGAGGGGGTCCTTCAGCACCTGCTCA
GACCCTTGTTCCAGAAGAGCAGCATGGGCCCCTTCTACTTGGGTTGCCAACTGATCTCCCTCAGGCCTGA
GAAGGATGGGGCAGCCACTGGTGTGGACACCACCTGCACCTACCACCCTGACCCTGTGGCCCCGGGCTG
GACATACAGCAGCTTTACTGGGAGCTGAGTCAGCTGACCCATGGTGTCACCCAACTGGGCTTCTATGTCC
TGGACAGGGATAGCCTCTTCATCAATGGTCACCACACTCTACAAAGGCAGTCAACTACATGACACATTCC
GCTTCTGCCTGGTCACCAACTTGACGATGGACTCCGTGTTGGTCACTGTCAAGGCATTGTTCTCCTCCAA
TTTGGACCCCAGCCTGGTGGAGCAAGTCTTTCTAGATAAGACCCTGAATGCCTCATTCCATTGGCTGGGC
TCCACCTACCAGTTGGTGGACATCCATGTGACAGAAATGGAGTCATCAGTTTATCAACCAACAAGCAGCT
CCAGCACCCAGCACTTCTACCTGAATTTCACCATCACCAACCTACCATATTCCCAGGACAAAGCCCAGCC
```

AGGCACCACCAATTACCAGAGGAACAAAAGGAATATTGAGGATGCGCTCAACCAACTCTTCCGAAACAGC
AGCATCAAGAGTTATTTTTCTGACTGTCAAGTTTCAACATTCAGGTCTGTCCCCAACAGGCACCACACCG
GGGTGGACTCCCTGTGTAACTTCTCGCCACTGGCTCGGAGAGTAGACAGAGTTGCCATCTATGAGGAATT
TCTGCGGATGACCCGCCGGGCAGGTACCCAGCTGCAGAACTTCACCCTGGACAGGAGCAGTGTCCTTGTG
GATGGGTATTCTCCCAACAGAAATGAGCCCTTAACTGGGAATTCTGACCTTCCCTTCTGGGCTGTCATCC
TCATCGGCTTGGCAGGACTCCTGGGACTCATCACATGCCTGATCTGCGGTGTCCTGGTGACCACCCGCCG
GCGGAAGAAGGAAGGAGAATACAACGTCCAGCAACAGTGCCCAGGCTACTACCAGTCACACCTAGACCTG
GAGGATCTGCAATGACTGGAACTTGCCGGTGCCTGGGGTGCCTTTCCCCCAGCCAGGGTCCAAAGAAGCT
TGGCTGGGGCAGAAATAAACCATATTGGTCGGTAAAAAAACTCTGAGCCT

>DEX0367_040.nt.3
GAAAAAAGAGGAAAGAAAGAAAAAGGGAGAATGWTRATACTATAGGCACTGGTTCTCTAATGCATGCTCG
AGCGGCGCAGTGTGATGGATCGCCGGGCAGGTGTGTCCTTGTGGATGGGTATTCTCCCAACAGAAATGAG
CCCTTAACTGGGAATTCTGACCTTCCCTTCTGGGCTGTCATCCTCATCGGCTTGGCAGGACTCCTGGGAC
TCATCACATGCCTGATCTGCGGTGTCCTGGTGACCACCCGCCGGCGGAAGAAGGAAGGAGAATACAACGT
CCAGCAACAGTGCCCAGGCTACTACCAGTCACACCTAGACCTGGAGGATCTGCAATGACTGGAACTTGCC
GGTGCCTGGGGTGCCTTTCCCCCAGCCAGGGTCCAAAGAAGCTTGGCTGGGGCAGAAATAAACCATATTG
GTCGGTAAAAAAACTCTGAGCCT

>DEX0367_040.nt.4
CATGCACCCTCCAGGTATTTCTTTCATTTCTAGGGTATTCTCCCAACAGAAATGAGCCCTTAACTGGGAA
TTCTGACCTTCCCTTCTGGGCTGTCATCCTCATCGGCTTGGCAGGACTCCTGGGACTCATCACATGCCTG
ATCTGCGGTGTCCTGGTGACCACCCGCCGGCGGAAGAAGGAAGGAGAATACAACGTCCAGCAACAGTGCC
CAGGCTACTACCAGTCACACCTAGACCTGGAGGATCTGCAATGACTGGAACTTGCCGGTGCCTGGGGTGC
CTTTCCCCCAGCCAGGGTCCAAAGAAGCTTGGCTGGGGCAGAAATAAACCATATTGGTCGGTAAAAAAAC
TCTGAGCCT

>DEX0367_041.nt.1
GCGGCGGCACAGCTCCCGCAGCTCCTGCTCTGGTCCGCCTCGGCCCGGACACCTCCCTCGCTCTCCCACA
CCACTGGCACCAGGCCCCGGACACCCGCTCTGCTGCAGGAGAATGGCTACTCATCACACGCTGTGGATGG
GACTGGCCCTGCTGGGGGTGCTGGGCGACCTGCAGGCAGCACCGGAGGCCCAGGTCTCCGTGCAGCCCAA
CTTCCAGCAGGACAAGTTCCTGGGGCGCTGGTTCAGCGCGGGCCTCGCCTCCAACTCGAGCTGGCTCCGG
GAGAAGAAGGCGGCGTTGTCCATGTGCAAGTCTGTGGTGGCCCCTGCCACGGATGGTGGCCTCAACCTGA
CCTCCACCTTCCTCAGGAAAAACCAGTGTGAGACCCGAACCATGCTGCTGCAGCCCGCGGGGTCCCTCGG
CTCCTACAGCTACCGGAGTCCCCACTGGGGCAGCACCTACTCCGTGTCAGTGGTGGAGACCGACTACGAC
CAGTACGCGCTGCTGTACAGCCAGGGCAGCAAGGGCCCTGGCGAGGACTTCCGCATGGCCACCCTCTACA
ATAAGTGCATGACGGAACAATAGGACTCCCCAGGGCTGAAGCTGGGATCCCGGCCAGCCAGGTGACCCCC
ACGCTCTGGATGTCTCTGCTCTGTTCCTTCCCCGAGCCCCTGCCCCGGCTCCCCGCCAAAGCACCCCTGC
CCACTCGGGCTTCATCCTGCACAATAAACTCCGGAAGCAAGTCAGTA

>DEX0367_041.nt.2
GCGGCGGCACAGCTCCCGCAGCTCCTGCTCTGGTCCGCCTCGGCCCGGACACCTCCCTCGCTCTCCCACA
CCACTGGCACCAGGCCCCGGACACCCGCTCTGCTGCAGGAGAATGGCTACTCATCACACGCTGTGGATGG
GACTGGCCCTGCTGGGGGTGCTGGGCGACCTGCAGGCAGCACCGGAGGCCCAGGTCTCCGTGCAGCCCAA
CTTCCAGCAGGACAAGTTCCTGGGGCGCTGGTTCAGCGCGGGCCTCGCCTCCAACTCGAGCTGGCTCCGG
GAGAAGAAGGCGGCGTTGTCCATGTGCAAGTCTGTGGTGGCCCCTGCCACGGATGGTGGCCTCAACCTGA
CCTCCACCTTCCTCAGGAAAAACCAGTGTGAGACCCGAACCATGCTGCTGCAGCCCGCGGGGTCCCTCGG
CTCCTACAGCTACCGGAGTCCCCGTGAGTGGGGCCTCCACCGGCCCCCTGGGCCCAGCCTGGGGGCGACA
CTTGCCGGGACGACTCTGGGCCAGCCCCCTGCCGCGGAGATCCATGGGGTGGGAGGTGACTGGGGCAGCA
CCTACTCCGTGTCAGTGGTGGAGACCGACTACGACCAGTACGCGCTGCTGTACAGCCAGGGCAGCAAGGG
CCCTGGCGAGGACTTCCGCATGGCCACCCTCTACAGCCGAACCCAGACCCCCAGGGCTGAGTTAAAGGAG
AAATTCACCGCCTTCTGCAAGGCCCAGGGCTTCACAGAGGATACCATTGTCTTCCTGCCCCAAACCGATA
AGTGCATGACGGAACAATAGGACTCCCCAGGGCTGAAGCTGGGATCCCGGCCAGCCAGGTGACCCCCACG
CTCTGGATGTCTCTGCTCTGTTCCTTCCCCGAGCCCCTGCCCCGGCTCCCCGCCAAAGCACCCCTGCCCA
CTCGGGCTTCATCCTGCACAATAAACTCCGGAAGCAAGTCAGTA

>DEX0367_041.nt.3
GCGGCGGCACAGCTCCCGCAGCTCCTGCTCTGGTCCGCCTCGGCCCGGACACCTCCCTCGCTCTCCCACA
CCACTGGCACCAGGCCCCGGACACCCGCTCTGCTGCAGGAGAATGGCTACTCATCACACGCTGTGGATGG
GACTGGCCCTGCTGGGGGTGCTGGGCGACCTGCAGGCAGCACCGGAGGCCCAGGTCTCCGTGCAGCCCAA

CTTCCAGCAGGACAAGTTCCTGGGGCGCTGGTTCAGCGCGGGCCTCGCCTCCAACTCGAGCTGGCTCCGG
GAGAAGAAGGCGGCGTTGTCCATGTGCAAGTCTGTGGTGGCCCCTGCCACGGATGGTGGCCTCAACCTGA
CCTCCACCTTCCTCAGGAAAAACCAGTGTGAGACCCGAACCATGCTGCTGCAGCCCGCGGGGTCCCTCGG
CTCCTACAGCTACCGGAGTCCCCACTGGGGCAGCACCTACTCCGTGTCAGTGGTGGAGACCGACTACGAC
CAGTACGCGCTGCTGTACAGCCAGGCCGAACCCAGACCCCCAGGGCTGAGTTAAAGGAGAAATTCACCGC
CTTCTGCAAGGCCCAGGGCTTCACAGAGGATACCATTGTCTTCCTGCCCCAAACCGATAAGTGCATGACG
GAACAATAGGACTCCCCAGGGCTGAAGCTGGATCCCGGCCAGCCAGGTGACCCCCACGCTCTGGATGTC
TCTGCTCTGTTCCTTCCCCGAGCCCCTGCCCCGGCTCCCCGCCAAAGCACCCCTGCCCACTCGGGCTTCA
TCCTGCACAATAAACTCCGGAAGCAAGTCAGTA

>DEX0367_041.nt.4
GCGGCGGCACAGCTCCCGCAGCTCCTGCTCTGGTCCGCCTCGGCCCGGACACCTCCCTCGCTCTCCCACA
CCACTGGCACCAGGCCCCGGACACCCGCTCTGCTGCAGGAGAATGGCTACTCATCACACGCTGTGGATGG
GACTGGCCCTGCTGGGGGTGCTGGGCGACCTGCAGGCAGCACCGGAGGCCCAGGTCTCCGTGCAGCCCAA
CTTCCAGCAGGACAAGGTGAGGGGCTTTCCTGCGTCATCCCCAAGGGCTACAGGACCCTGTCAGGGGAAG
GGCACTTTCCGGCTGGGTCTTCCCCCTGGGAGGAGTGAGCGAAGTCCCGCAGTGCCGGGCAGTGCCGGCC
AGGGACTGAGCGGGCGCGCAGGAAGGAGGCTGGGTAGCCGCCCGAGGAGGCTGCCCGCGAGAAGCCCTCC
CTGGGCACCGCGTCCCGTCTCCCCTGATGGGCCACGCAGACACCGGGCCACACACGCGCCCACGCCAGCC
AGACACGTCCACCCCTGTGGGCACGCGACCCCCAGAGGACACACGAGCGCACGTTCCACACCTGGGTGCC
AGGACACGGGCGGGTGGGGCACAGGGATGGCGACAAACACTCCGTGCGCGGTGGGGGTTGGGAGGGACGC
GCACCGCACAGACAGCCGGCGACGCGCGCTCTCGCCCGGGAGCTGCTCGGGGAAGCGCCGGAGCGCGGGT
CCCCGCGCCGCGCGCCCCTCCCTGGCGTCGAGGAGAACCCCAGCGGTCCGCAGAGCTGAGCCGAAGACCC
GCCCCGATCCCCGCCAGGAATGCGACGTCCTCTGCCGCCCGCTGTATGGGCCAGGCGCTCAGCCAGGGCA
CAGAATCGGGACCGGGGGCGGAGGGGCCGAAGCTGGCTGGTGGGCGGCGTGCGAGGTTCCTGGGGCGCTG
GTTCAGCGCGGGCCTCGCCTCCAACTCGAGCTGGCTCCGGGAGAAGAAGGCGGCGTTGTCCATGTGCAAG
TCTGTGGTGGCCCCTGCCACGGATGGTGGCCTCAACCTGACCTCCACCTTCCTCAGGAAAAACCAGTGTG
AGACCCGAACCATGCTGCTGCAGCCCGCGGGGTCCCTCGGCTCCTACAGCTACCGGAGTCCCCACTGGGG
CAGCACCTACTCCGTGTCAGTGGTGGAGACCGACTACGACCAGTACGCGCTGCTGTACAGCCAGGGCAGC
AAGGGCCCTGGCGAGGACTTCCGCATGGCCACCCTCTACAGCCGAACCCAGACCCCCAGGGCTGAGTTAA
AGGAGAAATTCACCGCCTTCTGCAAGGCCCAGGGCTTCACAGAGGATACCATTGTCTTCCTGCCCCAAAC
CGATAAGTGCATGACGGAACAATAGGACTCCCCAGGGCTGAAGCTGGATCCCGGCCAGCCAGGTGACCC
CCACGCTCTGGATGTCTCTGCTCTGTTCCTTCCCCGAGCCCCTGCCCCGGCTCCCCGCCAAAGCACCCCT
GCCCACTCGGGCTTCATCCTGCACAATAAACTCCGGAAGCAAGTCAGTA

>DEX0367_041.nt.5
GCGGCGGCACAGCTCCCGCAGCTCCTGCTCTGGTCCGCCTCGGCCCGGACACCTCCCTCGCTCTCCCACA
CCACTGGCACCAGGCCCCGGACACCCGCTCTGCTGCAGGAGAATGGCTACTCATCACACGCTGTGGATGG
GACTGGCCCTGCTGGGGGTGCTGGGCGACCTGCAGGCAGCACCGGAGGCCCAGGTCTCCGTGCAGCCCAA
CTTCCAGCAGGACAAGGTGAGGGGCTTTCCTGCGTCATCCCCAAGGGCTACAGGACCCTGTCAGGGGAAG
GGCACTTTCCGGCTGGGTCTTCCCCCTGGGAGGAGTGAGCGAAGTCCCGCAGTGCCGGGCAGTGCCGGCC
AGGGACTGAGCGGGCGCGCAGGAAGGAGGCTGGGTAGCCGCCCGAGGAGGCTGCCCGCGAGAAGCCCTCC
CTGGGCACCGCGTCCCGTCTCCCCTGATGGGCCACGCAGACACCGGGCCACACACGCGCCCACGCCAGCC
AGACACGTCCACCCCTGTGGGCACGCGACCCCCAGAGGACACACGAGCGCACGTTCCACACCTGGGTGCC
AGGACACGGGCGGGTGGGGCACAGGGATGGCGACAAACACTCCGTGCGCGGTGGGGGTTGGGAGGGACGC
GCACCGCACAGACAGCCGGCGACGCGCGCTCTCGCCCGGGAGCTGCTCGGGGAAGCGCCGGAGCGCGGGT
CCCCGCGCCGCGCGCCCCTCCCTGGCGTCGAGGAGAACCCCAGCGGTCCGCAGAGCTGAGCCGAAGACCC
GCCCCGATCCCCGCCAGGAATGCGACGTCCTCTGCCGCCCGCTGTATGGGCCAGGCGCTCAGCCAGGGCA
CAGAATCGGGACCGGGGGCGGAGGGGCCGAAGCTGGCTGGTGGGCGGCGTGCGAGGGTGAGGGACTGTCC
TCGGGAGGGGCGTGGCCAGACGGGGGTGCTGGGTGTCAGGGGCGTGGCCAGCTCCTGGGGCGGAGATGTG
AGGGGCGGGGTCATCTCCTGGGGCGGGGATTGAGGGGCGGGGGTCAGTTCCTGGGGCGGGGCGTGAGGGG
CGTGGCTGCTCGGGGGATTGGGCGTGGGGGCGGGGCCGGGTTGGAGACCGGAGGAGTAGACGGCAGAGGC
GCCCCCGCAGGTAGGCGCAGGTGGGTCCGGAGGGTCCTGGCCGACGCGGGTGGGGGTCGCTCGCCGCAGT
TTCCTGGGGCGCTGGTTCAGCGCGGGCCTCGCCTCCAACTCGAGCTGGCTCCGGGAGAAGAAGGCGGCGT
TGTCCATGTGCAAGTCTGTGGTGGCCCCTGCCACGGATGGTGGCCTCAACCTGACCTCCACCTTCCTCAG
GAAAAACCAGTGTGAGACCCGAACCATGCTGCTGCAGCCCGCGGGGTCCCTCGGCTCCTACAGCTACCGG
AGTCCCCGTGAGTGGGGCCTCCACCGGCCCCCTGGGCCCAGCCTGGGGGCGACACTTGCCGGGACGACTC
TGGGCCAGCCCCCTGCCGCGGAGATCCATGGGGTGGGAGGTGATGGCTGCCCCACCAGCGTCAGAGGCAA
AGGCCAGGCCTGGGCGTGACTACCCATGCACAAGTGTTAGGGACAGAGAGACCCTTCCTCCAGGGGGTTG
GATCCTCTCTGGAGCCCACCATTGTCTTGTCAGGCCCCTTCCCTGCCCTCTGGAGTTTTCCCCACATAAG
CAGCCCCCCAAGGCCCCTCCATATGCCTCCTCCCAATTCTCCTCCCCAGGACCCAGGGGTTTCCTCACTC

CCACCTGGGGAATGGCTCCCACGGGGAAACCTCTTCACTTCCGGTTCTGGCAGCGACTTCTGCGGCTGCA
CCAGGAATCCTGGTTTTCTGAGCCTGGCTCCCCCAGATTCTGGTTTGGGGACAGGGTTCACAGGCTGTGC
AGGCGAGAGCAGGGCACTGAGGTGCACCCCCTTCCCTGAAGCAGAGGTGAGGTTTGGGGGGCTGAGTCCCCGAC
TGCCCCCAAAGCCCACAGGTGCACCCCCTTCCCTGAAGCAGAGGTGAGGTTTGGGGGGCTGAGTCCCCGAC
5   AGGGTTGTCTCTTGGGTTCCCAGACTGGGGCAGCACCTACTCCGTGTCAGTGGTGGAGACCGACTACGAC
CAGTACGCGCTGCTGTACAGCCAGGGCAGCAAGGGCCCTGGCGAGGACTTCCGCATGGCCACCCTCTACA
GCCGAACCCAGACCCCCAGGGCTGAGTTAAAGGAGAAATTCACCGCCTTCTGCAAGGCCCAGGGCTTCAC
AGAGGATACCATTGTCTTCCTGCCCCAAACCGATAAGTGCATGACGGAACAATAGGACTCCCCAGGGCTG
AAGCTGGGATCCCGGCCAGCCAGGTGACCCCCACGCTCTGGATGTCTCTGCTCTGTTCCTTCCCCGAGCC
10   CCTGCCCCGGCTCCCCGCCAAAGCACCCCTGCCCACTCGGGCTTCATCCTGCACAATAAACTCCGGAAGC
AAGTCAGTA

>DEX0367_041.nt.6
CAATGGCAGCCGGCCCAATGGCAGCCGTAGCCGTGCGGGCCGCACGCGTTAGTCGCCCTGGCCGGCCTCG
15   TCACAGGGATACCCACCCACCATCCCAGGGTCTACAACATCCACAGCCGGACCGTCACCCGCTATCCTGC
CAACTCTATCGTGGTCGTAGGAGGCTGTCCTGTCTGCAGGGTTGGGGTGCTGGAGGACTGCTTCACCTTTC
CTGGGCATCTTCCTGGCCATCATCCTCTTCCGCATTGGGCCGGCTGCCATTGGGCAATGGCAGCCGCCCA
ATGGCAGCCGAACCCAGACCCCCAGGGCTGAGTTAAAGGAGAAATTCACCGCCTTCTGCAAGGCCCAGGG
CTTCACAGAGGATACCATTGTCTTCCTGCCCCAAACCGATAAGTGCATGACGGAACAATAGGACTCCCCA
20   GGGCTGAAGCTGGGATCCCGGCCAGCCAGGTGACCCCCACGCTCTGGATGTCTCTGCTCTGTTCCTTCCCC
CGAGCCCCTGCCCCGGCTCCCCGCCAAAGCACCCCTGCCCACTCGGGCTTCATCCTGCACAATAAACTCC
GGAAGCAAGTCAGTA

>DEX0367_041.nt.7
25   CCGCCCTGCTCGAGGCCTGGGCCAGGGACAGAGGGGGTGAGTGTTCAAGTCAGAACCTCCCTCCCGCAGCC
CCTTCATGAGGAGCCCCCTCCATGGGGGACATGGAGACCAGGCGCCCACTCTGTGCCAGGCCCCAGCAGC
TCCCAGGATGTGGGGCTTCAGCCTGGTGGGGGGGCATAGGGTAGAGGGTGCCCATGGGGGATACAGGGGAA
CCAACCATACTGGGCTCAGGCACAGCCTCCTAGGGGTGGACAGCCTGCTCTTGGCAGAGGTGGAGAAAGA
CCCTCTCTTTGTTTCCAGTGCCCAGGGAGAGGTGGGAGGTGATGGGGGATCTGTGCAGTTTGGGGGGCTCA
30   GTCAAGACGAGCTCTGCGTTACGGGAGGAACAGGAAGCCCAGTGGGAAAATTGGCCTAAGTCTGGGGTTC
TGACGACAGCCCCTGGCTTCTTTCTTGGCAGATAAGTGCATGACGGAACAATAGGACTCCCCAGGGCTGA
AGCTGGGATCCCGGCCAGCCAGGTGACCCCCACGCTCTGGATGTCTCTGCTCTGTTCCTTCCCCGAGCCC
CTGCCCCGGCTCCCCGCCAAAGCACCCCTGCCCACTCGGGCTTCATCCTGCACAATAAACTCCGGAAGCA
AGTCAGTA

35
>DEX0367_042.nt.1
CGGGTGGGGAAGGCAGGAGGCGGCGATCCCGGCGGAGGGGGCCGTTCGCCAGCTCTGAGGCAGAAAGTGC
CACGACTCCACACGCGCGCACGCAGCCAGCGAGCGGCCGGAGCGGACGGCAGACGGGGCGGGCGGCGTCA
GGGTCGCAGCCGTCTACAGCTGCTCGGGGGCCGGTTTCTTGGCGGAGGCTTGGCCGGCTCCTCTCTCCCGGC
40   TCCGCGGCGGCTGCGAAGGCGGCGGCTCCTGCCCTCTCGCTTTCCCTCTCGCGTCTCTGGCTGCAGGTGA
AAGGAAAGCAAGCCAGGATGGATATTTACGACACTCAAACCTTGGGGGTTGTGGTCTTTGGAGGATTCAT
GGTTGTTTCTGCCATTGGCATCTTCCTGGTGTCGACTTTCTCCATGAAGGAAACGTCATATGAAGAAGCC
CTAGCCAACCAGCGCAAGGAGATGGCGAAAACTCACCACCAGAAAGTCGAGAAGAAAAAGAAGGAGAAAA
CAGTGGAGAAGAAAGGAAAGACCAAGAAAAAGGAAGAGAAACCTAATGGGAAGATACCTGATCATGATCC
45   AGCCCCCAATGTGACTGTCCTCCTTCGAGAACCAGTGCGGGCTCCTGCTGTGGCTCCAACCCCA
GTGCAGCCCCCCATTATCGTTGCTCCTGTCGCCACAGTTCCAGCCATGCCCCAGGAGAAGCTGGCCTCCT
CCCCCAAGGACAAAAAGAAGAAGGAGAAAAAAGTGGCAAAAGTGGAACCAGCTGTCAGCTCTGTAGTGAA
TTCCATCCAGGTTCTCACTTCGAAGGCTGCCATCTTGGAAACTGCTCCCAAGGAGGTGCCGATGGTGGTG
GTGCCCCCAGTGGGTGCCAAGGGCAACACACCAGCCACTGGCACTACTCAGGGCAAAAAGGCGGAGGGGA
50   CTCAGAATCAAAGCAAAAAGGCTGAAGGAGCCCCAAACCAGGGCAGAAAGGCAGAGGGAACCCCAAACCA
GGGCAAAAAGACAGAGGGAACCCCAAACCAAGGGAAAAAGGCAGAGGGAACCCCAAACCAAGGCAAAAAG
GCAGAAGGAACCCCAAACCAAGGCAAAAAGGCGGAGGGGGCCCAGAACCAGGGTAAAAAGGTAGATACAA
CCCCAAACCAGGGGAAAAAGGTGGAGGGGGCCCCAACCCAGGGCAGAAAGGCCGAGGGGGCTCAGAACCA
GGCCAAAAAGGTAGAAGGGGCCCAGAACCAGGGCAAAAAGGCAGAGGGGGCCCAGAATCAGGGCAAAAAG
55   GGAGAGGGGGCCCAGAACCAGGGCAAGAAGGCTGAGGGGGCCCAGAACCAGGGCAAGAAGGCCGAGGGGG
CTCAGAACCAGGGCAAAAAGGCCGAGGGAGCCCAGAACCAGGGCCAAAAAGGAGAGGGAGCCCAGAATCA
GGGTAAAAAGACAGAAGGGGCTCAGGGCAAAAAGGCAGAAAGGAGTCCCAACCAAGGCAAAAAAGGAGAG
GGAGCTCCCATCCAGGGCAAAAAGGCAGATTCGGTTGCTAATCAGGGCACAAAGGTAGAGGGTATTACAA
ACCAGGGGAAAAAGCAGAAGGGTCCCCCAGTGAAGGCAAAAAGGCAGAAGGGTCCCCAACCAAGGCAA
60   AAAGGCAGACGCAGCTGCCAATCAGGGTAAAAAGACAGAGTCAGCTTCTGTCCAGGGCAGAAATACAGAT

```
GTGGCCCAGAGCCCAGAGGCACCAAAGCAAGAGGCTCCTGCCAAGAAGAAGTCTGGTTCAAAGAAAAAAG
GTGAGCCTGGGCCCCCAGATGCCGACGGCCCTCTCTACCTCCCCTACAAGACGCTGGTCTCCACGGTTGG
GAGCATGGTGTTCAACGAGGGCGAGGCCCAGCGGCTCATCGAGATCCTGTCTGAGAAGGCTGGCATCATT
CAGGACACCTGGCACAAGGCCACTCAGAAGGGTGACCCTGTGGCGATTCTGAAACGCCAGCTGGAAGAGA
AGGAAAAACTGCTGGCCACAGAACAGGAAGATGCGGCTGTCGCCAAGAGCAAACTGAGGGAGCTCAACAA
GGAGATGGCAGCAGAAAAGGCCAAAGCAGCAGCCGGGGAGGCCAAAGTGAAAAAGCAGCTGGTGGCCCGG
GAGCAGGAGATCACGGCTGTGCAGGCACGCATGCAGGCCAGCTACCGGGAGCACGTGAAGGAGGTGCAGC
AGCTGCAGGGCAAGATCCGGACTCTTCAGGAGCAGCTGGAGAATGGCCCCAACACGCAGCTGGCCCGCCT
GCAGCAGGAGAACTCCATCCTGCGGGATGCCTTGAACCAGGCCACGAGCCAGGTGGAGAGCAAGCAGAAC
GCAGAGCTGGCCAAGCTTCGGCAGGAGCTCAGCAAGGTCAGCAAAGAGCTGGTGGAGAAGTCAGAGGCTG
TGCGGCAAGATGAGCAGCAGCGGAAAGCTCTGGAAGCCAAGGCAGCTGCCTTCGAGAAGCAGGTCCTGCA
GCTGCAGGCGTCCCACAGGGAGAGTGAGGAGGCCCTGCAGAAGCGCCTGGACGAGGTCAGCCGGGGAGCTG
TGCCACACGCAGAGCAGCCACGCCAGCCTCCGGGCGGATGCCGAGAAGGCCCAGGAGCAACAGCAGCAGA
TGGCCGAGCTGCACAGCAAGTTACAGTCCTCCGAGGCGGAGGTGCGCAGCAAATGCGAGGAGCTGAGTGG
CCTCCACGGGCAGCTCCAGGAGGCCAGGGCGGAGAACTCCCAGCTCACAGAGAGAATCCGTTCCATTGAG
GCCCTGCTGGAGGCGGGCCAGGCGCGGGATGCCCAGGACGTCCAGGCCAGCCAGGCGGAGGCTGACCAGC
AGCAGACTCGCCTCAAGGAGCTGGAGTCCCAGGTGTCGGGTCTGGAGAAGGAGGCCATCGAGCTCAGGGA
GGCCGTCGAGCAGCAGAAAGTGAAGAACAATGACCTCCGGGAGAAGAACTGGAAGGCCATGGAGGCACTG
GCCACGGCCGAGCAGGCCTGCAAGGAGAAGCTGCACTCCCTGACCCAGGCCAAGGAGGAATCGGAGAAGC
AGCTCTGTCTGATTGAGGCGCAGACCATGGAGGCCCTGCTGGCTCTGCTCCCAGAACTCTCTGTCTTGGC
ACAACAGAATTACACCGAGTGGCTGCAGGATCTCAAAGAGAAAGGCCCCACGCTGCTGAAGCACCCGCCA
GCTCCCGCGGAGCCCTCCTCGGACCTGGCCTCCAAGTTGAGGGAGGCCGAGGAGACGCAGAGCACACTGC
.AGGCCGAGTGTGACCAGTACCGCAGCATCCTGGCGGGACGAGGGGCATGCTCAGAGACCTGCAGAAGAG
CGTGGAGGAGGAGGAGCAGGTGTGGAGGGCCAAGGTGGGCGCCGCAGAGGAGGAGCTCCAGAAGGTATAT
GCAGCCCTGCCTGCCTCCCCGGCGCCGTGGAGCACATGAAGCCCAGGGGAGAGTTGACAGGATTGTCTCTG
TCACCTGTGAGACTCGTGTGTGTCTCGGTGGGGCTTTTCCACGACACCCTATGAAGCACAGTTGCGGGATGA
CAGCTCAGTCCGTCCTCCAGGAGGCCCCCGCCTAGGAAGACACTGCCAGATGGCGCCAGGGTGACCGTCC
CAGCTGGGCGGAGTGGGCACCCCCGGGGCTTCCTAAGTTGGGTCAGTCCCGCTCGGGTCCTAGCATGGCC
TCGGAAGGAGACACCCCCAGGTGCTCGTGTCTGACGGAGCCTTCCCCTTGCTTTGTAGTCCCGGGTCACA
GTGAAGCATCTCGAAGAGATTGTAGAGAAGCTAAAAGGAGAACTTGAAAGTTCGGACCAGGTATGTGCTC
TGCTCAGAGAACTCTTTCTCTTGTCCCGGGAGCCTCTTTGAAAGTGACGTTAATGCAGAGGACACTTCAC
TGATTTCTTCCTCATGGACCGTGTGCCTTAAGAGAGGCATTCTCTCTCAAGCCAAGGGATCTTGTCCATC
CTGTTTTTCCTTAATGAGAGACATTGACTACCTCAGAGGACCCTGTCACAGTCAAGGTGGCCCCACTGCCCCCTT
CAGACGCCACAGGGAGCCCTGGAGAAGGTACCCTTGGTACCTGCTTTTGCAAGTAAATCTAAGAGCATGT
CAGTCCCCAGGAGTGCCCGGGTGCAGCCTGCCACTTTCCGTAAGGGGTTCTCTTGCCGCGGACTAGTTTC
TCGTTTGCTTGGTAGCTGGTGGGCCTCAGGGATCTCCCGAGAGGCCTCACCTATCCCAGGGCCAGCCTGC
TCCCTGAGGTGAGTGCCCTTTGAGCAGAGGCTGCTGGCGCATGCGTGTGCGTGTCTGTCATCTGTCCGT
GAGGGAGCACACGTTGCATTTGGAGGCAGAGCTGGAAAAGCACATGGCGGCCGCCAGCGCCGAGTGCCAG
AACTACGCCAAGGAGGTGGCAGGGCTGAGGCAACTTCTCCTAGAATCTCAATCTCAGCTCGATGCCGCCA
AGAGCGAAGCCCAGAAACAGAGCGATGAGCTTGCCCTGGTCAGGCAGCAGTTGAGTGAAATGAAGAGCCA
CGTAGAGGATGGTGACATAGCTGGGGCCCCAGCTTCCTCCCCAGAGGCGCCCCCAGCCGAGCAGGACCCC
GTTCAGCTGAAGACGCAGCTGGAGTGGACAGAAGCCATCCTGGAGGATGAGCAGACACAGCGGCAGAAGC
TCACGGCCGAGTTTGAGGAGGCTCAGACCTCGGCATGTCGGTTACAAGAAGAATTGGAGAAGCTCCGCAC
AGCCGGCCCCCTAGAGTCTTCAGAAACAGAGCAGCAGCCGCCGACGAGACTCACAGCTGAAGGAGAGACTAGAAAAAGAGAAG
AAGTTAACAAGTGACCTGGGGCGCGCCGCCACGAGACTCAGGAGCTTCTGAAGACGACCCAGGAGCAGC
TGGCAAGGGAGAAGGACACGGTGAAGAAGCTGCAGGAACAGCTGGAAAAGGCAGAGGACGGCAGCAGCTC
AAAGGAGGGCACCTCTGTCTGAGTTTCCTCTTTGGAAAAAGAAGTTACTGTTCAACTTACCAAAATGCCT
TACACATTCCTTACAAATAAACCAACCAACCTACACAGCGTTATCCAGGCCCAACTTCCGGTACTCCGAG
AGAAGCCATGAGAGACAAGTCTCTTAGAGCCACAGAAGTAGACCTTCCAGAGCCCCAGTTTGTAAATGAA
CCTGTGTCACATTTGATAAACACTATCCTGGGCGCAGCCCCGGGCCACCGCCGAGTGACGCCAAAGCCCT
GGTTGACTCTGACAGCCCCGTGGGTGTGTGGGAGGCCGGGCGCTCTGGGTCTGTCTGTCAGTGCAATCG
TTTAGTGTTTTTTCAGTGGGGCGGGGCGGGAAGCGGGTGGGACCGGGCAGCCAGTTCTCAAAGGCTGTGG
GGCCGACTGGAGGCCACAGCCCCTCACCCCTAGACGTTGCCAACCAGAACTGACGTGTGACCTCCTGGGT
GTTGATGCCATTAAAACCAACGTTGGTGCCCGGTG
```

```
>DEX0367_042.nt.2
CGGGTGGGGAAGGCAGGAGGCGGCGATCCCGGCGGAGGGGGCCGTTCGCCAGCTCTGAGGCAGAAAGTGC
CACGACTCCACACGCGCGCACGCAGCCAGCGAGCGGCCGGAGCGGACGGCAGACGGGGCGGGCGGCGTCA
GGGTCGCAGCGTCTACAGCTGCTCGGGGGCGGTTTCTTGGCGGAGGCTTGGCCGGCTCCTCTCTCCCGGC
```

```
TCCGCGGCGGCTGCGAAGGCGGCGGCTCCTGCCCTCTCGCTTTCCCTCTCGCGTCTCTGGCTGCAGGTGA
AAGGAAAGCAAGCCAGGATGGATATTTACGACACTCAAACCTTGGGGGTTGTGGTCTTTGGAGGATTCAT
GGTTGTTTCTGCCATTGGCATCTTCCTGGTGTCGACTTTCTCCATGAAGGAAACGTCATATGAAGAAGCC
CTAGCCAACCAGCGCAAGGAGATGGCGAAAACTCACCACCAGAAAGTCGAGAAGAAAAAGAAGGAGAAAA
CAGTGGAGAAGAAAGGAAAGACCAAGAAAAAGGAAGAGAAACCTAATGGGAAGATACCTGATCATGATCC
AGCCCCCAATGTGACTGTCCTCCTTCGAGAACCAGTGCGGGCTCCTGCTGTGGCTGTGGCTCCAACCCCA
GTGCAGCCCCCATTATCGTTGCTCCTGTCGCCACAGTTCCAGCCATGCCCCAGGAGAAGCTGGCCTCCT
CCCCCAAGGACAAAAAGAAGAAGGAGAAAAAAGTGGCAAAAGTGGAACCAGCTGTCAGCTCTGTAGTGAA
TTCCATCCAGGTTCTCACTTCGAAGGCTGCCATCTTGGAAACTGCTCCCAAGGAGGTGCCGATGGTGGTG
GTGCCCCCAGTGGGTGCCAAGGGCAACACACCAGCCACTGGCACTACTCAGGGCAAAAAGGCGGAGGGGA
CTCAGAATCAAAGCAAAAAGGCTGAAGGAGCCCCAAACCAGGGCAGAAAGGCAGAGGGAACCCCAAACCA
GGGCAAAAAGACAGAGGGAACCCCAAACCAAGGGAAAAAGGCAGAGGGAACCCCAAACCAAGGCAAAAAG
GCAGAAGGAACCCCAAACCAAGGCAAAAAGGCGGAGGGGGCCCAGAACCAGGGTAAAAAGGTAGATACAA
CCCCAAACCAGGGGAAAAAGGTGGAGGGGGCCCCAACCCAGGGCAGAAGGGCCGAGGGGGCTCAGAACCA
GGCCAAAAAGGTAGAAGGGGCCCAGAACCAGGGCAAAAAGGCAGAGGGGGCCCAGAATCAGGGCAAAAAG
GGAGAGGGGGCCCAGAACCAGGGCAAGAAGGCTGAGGGGGCCCAGAACCAGGGCAAGAAGGCCGAGGGGG
CTCAGAACCAGGGCAAAAAGGCCGAGGGAGCCCAGAACCAGGGCCAAAAAGGAGAGGGAGCCCAGAATCA
GGGTAAAAAGACAGAAGGGGCTCAGGGCAAAAAGGCAGAAAGGAGTCCCAACCAAGGCAAAAAAGGAGAG
GGAGCTCCCATCCAGGGCAAAAAGGCAGATTCGGTTGCTAATCAGGGCACAAAGGTAGAGGGTATTACAA
ACCAGGGGAAAAAGCAGAAGGGTCCCCCAGTGAAGGCAAAAAGGCAGAAGGGTCCCCCAACCAAGGCAA
AAAGGCAGACGCAGCTGCCAATCAGGGTAAAAAGACAGAGTCAGCTTCTGTCCAGGGCAGAAATACAGAT
GTGGCCCAGAGCCCAGAGGCACCAAAGCAAGAGGCTCCTGCCAAGAAGAAGTCTGGTTCAAAGAAAAAAG
GTGAGCCTGGGCCCCCAGATGCCGACGGCCCTCTCTACCTCCCCTACAAGACGCTGGTCTCCACGGTTGG
GAGCATGGTGTTCAACGAGGGCGAGGCCCAGCGGCTCATCGAGATCCTGTCTGAGAAGGCTGGCATCATT
CAGGACACCTGGCACAAGGCCACTCAGAAGGGTGACCCTGTGGCGATTCTGAAACGCCAGCTGGAAGAGA
AGGAAAAACTGCTGGCCACAGACAGGAAGATGCGGGCTGTCGCCAAGAGCAAACTGAGGGAGCTCAACAA
GGAGATGGCAGCAGAAAAGGCCAAAGCAGCAGCCGGGGAGGCCAAAGTGAAAAAGCAGCTGGTGGCCCGG
GAGCAGGAGATCACGGCTGTGCAGGCACGCATGCAGGCCAGCTACCGGGAGCACGTGAAGGAGGTGCAGC
AGCTGCAGGGCAAGATCCGGACTCTTCAGGAGCAGCTGGAGAATGGCCCCAACACGCAGCTGGCCCGCCT
GCAGCAGGAGAACTCCATCCTGCGGGATGCCTTGAACCAGGCCACGAGCCAGGTGGAGAGCAAGCAGAAC
GCAGAGCTGGCCAAGCTTCGGCAGGAGCTCAGCAAGGTCAGCAAAGAGCTGGTGGAGAAGTCAGAGGCTG
TGCGGCAAGATGAGCAGCAGCGGAAAGCTCTGGAAGCCAAGGCAGCTGCCTTCGAGAAGCAGGTCCTGCA
GCTGCAGGCGTCCCACAGGGAGAGTGAGGAGGCCCTGCAGAAGCGCCTGGACGAGGTCAGCCGGGAGCTG
TGCCACACGCAGAGCAGCCACGCCAGCCTCCGGGCGGATGCCGAGAAGGCCCAGGAGCAACAGCAGCAGA
TGGCCGAGCTGCACAGCAAGTTACAGTCCTCCGAGGCGGAGGTGCCGCAGCAAATGCGAGGAGCTGAGTGG
CCTCCACGGGCAGCTCCAGGAGGCCAGGCGGCGGAGAATCCCAGCTCACAGAGAGAATCCGTTCCATTGAG
GCCCTGCTGGAGGCGGGCCAGGCGCGGGATGCCCAGGACGTCCAGGCCAGCCAGGCGGAGGCTGACCAGC
AGCAGACTCGCCTCAAGGAGCTGGAGTCCCAGGTGTCGGGTCTGGAGAAGGAGGCCATCGAGCTCAGGGA
GGCCGTCGAGCAGCAGAAAGTGAAGAACAATGACCTCCGGGAGAAGAACTGGAAGGCCATGGAGGCACTG
GCCACGGCCGAGCAGGCCTGCAAGGAGAAGCTGCACTCCCTGACCCAGGCCAAGGAGGAATCGGAGAAGC
AGCTCTGTCTGATTGAGGCGCAGACCATGGAGGCCCTGCTGGCTCTGCTCCCAGAACTCTCTGTCTTGGC
ACAACAGAATTACACCGAGTGGCTGCAGGATCTCAAAGAGAAAGGCCCCACGCTGCTGAAGCACCCGCCA
GCTCCCGCGGAGCCCTCCTCGGACCTGGCCTCCAAGTTGAGGGAGGCCGAGGAGACGCAGAGCACACTGC
AGGCCGAGTGTGACCAGTACCGCAGCATCCTGGCGGAGACGGAGGGCATGCTCAGAGACCTGCAGAAGAG
CGTGGAGGAGGAGGAGCAGGTGTGGAGGGCCAAGGTGGGCGCCGCAGAGGAGGAGCTCCAGAAGGTATAT
GCAGCCCTGCCTGCCTCCCGGCGCCGTGGAGCACATGAAGCCCAGGGGAGAGTTGACAGGATTGTCTCTG
TCACCTGTGAGACTCGTGTGTCTCGGTGGGGCTTTTCCACGACACCCTATGAAGCACACAGTTGCGGGATGA
CAGCTCAGTCCGTCCTCCAGGAGGCCCCCGCCTAGGAAGACACTGCCAGATGGCGCCAGGGTGACCGTCC
CAGCTGGGCGGAGTGGGCGCACCCCCGGGGCTTCCTAAGTTGGGTCAGTCCCGCTCGGGTCCTAGCATGGCC
TCGGAAGGAGACACCCCCAGGTGCTCGTGTCTGACGGAGCCTTCCCCTTGCTTTGTAGTCCCGGGTCACA
GTGAAGCATCTCGAAGAGATTGTAGAGAAGCTAAAAGGAGAACTTGAAAGTTCGGACCAGGTGAGGGAGC
ACACGTTGCATTTGGAGGCAGAGCTGGAAAAGCACATGGCGGCCGCCAGCGCCGAGTGCCAGAACTACGC
CAAGGAGGTGGCAGGGCTGAGGCAACTTCTCCTAGAATCTCAATCTCAGCTCGATGCCGCCAAGAGCGAA
GCCCAGAAACAGAGCGATGAGCTTGCCCTGGTCAGGCAGCAGTTGAGTGAAATGAAGAGCACGTAGAGG
ATGGTGACATAGCTGGGGCCCCAGCTTCCTCCCCAGAGGCGCCCCCAGCCGAGCAGGACCCCGTTCAGCT
GAAGACGCAGCTGGAGTGGACAGAAGCCATCCTGGAGGATGAGCAGACACAGCGGCAGAAGCTCACGGCC
GAGTTTGAGGAGGCTCAGACCTCGGCATGTCGGTTACAAGAAGAATTGGAGAAGCTCCGCACAGCCGGCC
CCCTAGAGTCTTCAGAAACAGAGGAGGCCTCACAGCTGAAGGAGAGACTAGAAAAAGAGAAGAAGTTAAC
AAGTGACCTGGGGCGCGCCGCCACGAGACTGCAGGAGCTTCTGAAGACGACCCAGGAGCAGCTGGCAAGG
GAGAAGGACACGGTGAAGAAGCTGCAGGAACAGCTGGAAAAAGGCAGAGGACGGCAGCAGCTCAAAGGAGG
```

GCACCTCTGTCTGAGTTTCCTCTTTGGAAAAAGAAGTTACTGTTCAACTTACCAAAATGCCTTACACATT
CCTTACAAATAAACCAACCAACCTACACAGCAGCGTTATCCAGGCCCAACTTCCGGTACTCCGAGAGAAGCCA
TGAGAGACAAGTCTCTTAGAGCCACAGAAGTAGACCTTCCAGAGCCCCAGTTTGTAAATGAACCTGTGTC
ACATTTGATAAACACTATCCTGGGCGCAGCCCCGGGCCACCGCCGAGTGACGCCAAAGCCCTGGTTGACT

5   CTGACAGCCCCGTGGGTGTGTGGGAGGCCGGGCGCTCTGGGGTCTGTCTGTCAGTGCAATCGTTTAGTGT
TTTTTCAGTGGGGCGGGGCGGGAAGCGGGTGGGACCGGGCAGCCAGTTCTCAAAGGCTGTGGGGCCGACT
GGAGGCCACAGCCCCTCACCCCTAGACGTTGCCAACCAGAACTGACGTGTGACCTCCTGGGTGTTGATGC
CATTAAAACCAACGTTGGTGCCCGGTG

10   >DEX0367_042.nt.3
CGGGTGGGGAAGGCAGGAGGCGGCGATCCCGGCGGAGGGGGCCGTTCGCCAGCTCTGAGGCAGAAAGTGC
CACGACTCCACACGCGCGCACGCAGCCAGCGAGCGGCCGGAGCGGACGGCAGACGGGGCGGGCGGCGTCA
GGGTCGCAGCGTCTACAGCTGCTCGGGGGCGGTTTCTTGGCGGAGGCTTGGCCGGCTCCTCTCTCCCGGC
TCCGCGGCGGCTGCGAAGGCGGCGGCTCCTGCCCTCTCGCTTTCCCTCTCGCGTCTCTGGCTGCAGGTGA

15   AAGGAAAGCAAGCCAGGATGGATATTTACGACACTCAAACCTTGGGGGTTGTGGTCTTTGGAGGATTCAT
GGTTGTTTCTGCCATTGGCATCTTCCTGGTGTCGACTTTCTCCATGAAGGAAACGTCATATGAAGAAGCC
CTAGCCAACCAGCGCAAGGAGATGGCGAAAACTCACCACCAGAAAGTCGAGAAGAAAAAGAAGGAGAAAA
CAGTGGAGAAGAAAGGAAAGACCAAGAAAAAGGAAGAGAAACCTAATGGGAAGATACCTGATCATGATCC
AGCCCCCAATGTGACTGTCCTCCTTCGAGAACCAGTGCGGGCTCCTGCTGTGGCTGTGGCTCCAACCCCA

20   GTGCAGCCCCCCATTATCGTTGCTCCTGTCGCCACAGTTCCAGCCATGCCCCAGGAGAAGCTGGCCTCCT
CCCCCAAGGACAAAAAGAAGAAGGAGAAAAAAGTGGCAAAAGTGGAACCAGCTGTCAGCTCTGTAGTGAA
TTCCATCCAGGTTCTCACTTCGAAGGCTGCCATCTTGGAAACTGCTCCCAAGGAGGTGCCGATGGTGGTG
GTGCCCCCAGTGGGTGCCAAGGGCAACACACCAGCCACTGGCACTACTCAGGGCAAAAAGGCGGGAGGGGA
CTCAGAATCAAAGCAAAAAGGCTGAAGGAGCCCCAAACCAGGGCAGAAAGGCAGAGGGAACCCCAAACCA

25   GGGCAAAAAGACAGAGGGAACCCCAAACCAAGGGCAGAGGGAACCCCAAACCAAGGCAAAAAG
GCAGAAGGAACCCCAAACCAAGGGCAAAAAGGCGGAGGGGGGCCCAGAACCAGGGTAAAAAGGTAGATACAA
CCCCAAACCAGGGGGAAAAAGGTGGAGGGGGCCCCAACCCAGGGCAGAAAGGCCGAGGGGGCTCAGAACCA
GGCCAAAAAGGTAGAAGGGGCCCAGAACCAGGGCAAAAAGGCAGAGGGGGCCCAGAATCAGGGCAAAAAG
GGAGAGGGGGCCCAGAACCAGGGCAAGAAGGCTGAGGGGGCCCAGAACCAGCCCCCAATGTGACTGTCCT

30   CCTTCGAGAACCAGTGCGGGCTCCTGCTGTGGCTGTGGCTCCAACCCCAGTGCAGCCCCCCATTATCGTT
GCTCCTGTCGCCACAGTTCCAGCCATGCCCCAGGAGAAGCTGGCCTCCTCCCCCAAGGACAAAAAGAAGA
AGGAGAAAAACGTGGCCAAAGTGGAACCAGCTGTCAGCTCTGTAGTGAATTCCATCCAGGTTCTCACTTC
GAAGGCTGCCAGGTTGGGAGCATGGTGTTCAACGAGGGCGAGGCCCAGCGGCTCATCGAGATCCTGTCTG
AGAAGGCTGGCATCATTCAGGACACCTGGCACAAGGCCACTCAGAAGGGTGACCCTGTGGCGATTCTGAA

35   ACGCCAGCTGGAAGAGAAGGAAAAACTGCTGGCCACAGAACAGGAAGATGCGGCTGTCGCCAAGAGCAAA
CTGAGGGAGCTCAACAAGGAGATGGCAGCAGAAAAGGCCAAAGCAGCAGCCGGGGAGGCCAAAGTGAAAA
AGCAGCTGGTGGCCCGGGAGCAGGAGATCACGGCTGTGAGGGAGCACACGTTGCATTTGGAGGCAGAGCT
GGAAAAGCACATGGCGGCCGCCAGCGCCGAGTGCCAGAACTACGCCAAGGAGGTGGCAGGGCTGAGGCAA
CTTCTCCTAGAATCTCAATCTCAGCTCGATGCCGCCAAGAGCGAAGCCCAGAAACAGAGCGATGAGCTTG

40   CCCTGGTCAGGCAGCAGTTGAGTGAAATGAAGAGCCACGTAGAGGATGGTGACATAGCTGGGGCCCCAGC
TTCCTCCCCAGAGGCGCCCCCAGCCGAGCAGGACCCCGTTCAGCTGAAGACGCAGCTGGAGTGGACAGAA
GCCATCCTGGAGGATGAGCAGACACAGCGGCAGAAGCTCACGGCCGAGTTTGAGGAGGCTCAGACCTCGG
CATGTCGGTTACAAGAAGAATTGGAGAAGCTCCGCACAGCCGGCCCCCTAGAGTCTTCAGAAACAGAGGA
GGCCTCACAGCTGAAGGAGAGACTAGAAAAAGAGAAGAAGTTAACAAGTGACCTGGGGCGCGCCGCCACG

45   AGACTGCAGGAGCTTCTGAAGACGACCCAGGAGCAGCTGGCAAGGGAGAAGGACACGGTGAAGAAGCTGC
AGGAACAGCTGGAAAAGGCAGAGGACGGCAGCAGCTCAAAGGAGGGCACCTCTGTCTGAGTTTCCTCTTT
GGAAAAAGAAGTTACTGTTCAACTTACCAAAATGCCTTACACATTCCTTACAAATAAACCAACCAACCTA
CACAGCGTTATCCAGGCCCAACTTCCGGTACTCCGAGAGAAGCCATGAGAGACAAGTCTCTTAGAGCCAC
AGAAGTAGACCTTCCAGAGCCCCAGTTTGTAAATGAACCTGTGTCACATTTGATAAACACTATCCTGGGC

50   GCAGCCCCGGGCCACCGCCGAGTGACGCCAAAGCCCTGGTTGACTCTGACAGCCCCGTGGGTGTGTGGGA
GGCCGGGCGCTCTGGGGTCTGTCTGTCAGTGCAATCGTTTAGTGTTTTTTCAGTGGGGCGGGGCGGGAAG
CGGGTGGGACCGGGCAGCCAGTTCTCAAAGGCTGTGGGGCCGACTGGAGGCCACAGCCCCTCACCCCTAG
ACGTTGCCAACCAGAACTGACGTGTGACCTCCTGGGTGTTGATGCCATTAAAACCAACGTTGGTGCCCGG
TG

55   >DEX0367_043.nt.1
GGGCAGGTCTCAGCGCTCCTCCCCCTGCTCCGCTCCTCTGCAGGGCCCAGGCGCCCTTGGCCTTAGGACC
CAACTTCTCTTACCGCCATGGAGTTCGACCTGGGAGCAGGTGAGCTCCTGGGGAGTGTGGACTGGAGGTG
CAGGGGGCCGGACTCAAGCCCAGAAGCTGCCTGCACCAACCACCCAACTTCTCTCCACAGCCCTGGAGCC

60   CACCTCCCAGAAGCCCGGTGTGGGGGCGGGCCACGGGGGAGATCCCAAGCTCAGTCCCCACAAAGTTCAG

GGCCGGTCGGAGGCAGGGGCAGGTCCGGGTCCAAAGCAAGGACACCACAGCTCTTCCGACTCCAGCAGCA
GCTCCAGCGATTCGGACACGGATGTGAAGTCCCACGCTGCTGGCTCCAAGCAGCACGAGAGCATCCCGGG
CAAGGCCAAGAAGCCCAAAGTGAAGAAGAAGGAGAAGGGCAAGAAGGAGAAGGGCAAGAAGAAGGAGGCT
CCCCACTGAAGGGCCCTGGACAGGGCTCATTAAACCTTCCTCTCTGCCTTC

5 >DEX0367_043.nt.2
GGGCAGGTCTCAGCGCTCCTCCCCCTGCTCCGCTCCTCTGCAGGGCCCAGGCGCCCTTGGCCTTAGGACC
CAACTTCTCTTACCGCCATGGAGTTCGACCTGGGAGCAGGTGAGCTCCTGGGGAGTGTGGACTGGAGGTG
CAGGGGGCCGGACTCAAGCCCAGAAGCTGCCTGCACCAACCACCCAACTTCTCTCCACAGCCCTGGAGCC
10 CACCTCCCAGAAGCCCGGTGTGGGGGCGGGCCACGGGGGAGATCCCAAGCTCAGTCCCCACAAAGTTCAG
GGCCGGTCGGAGGCAGGGGCAGGTCCGGGTCCAAAGGTAAGTCGCCTCATCACCGGCTGCGGAGGGGCGG
GAAGGCTGGGGTTGCCCCTGACCCCAGGGTCCTGCCTTAGGCCTCCAACTTCAGGGGGCTGGGTAAGGGG
CGCCGCCTCACTGCCGCACCTCCATCCAGCAAGGACACCACAGCTCTTCCGACTCCAGCAGCAGCTCCAG
CGATTCGGACACGGATGTGAAGGTAAGGGGCTCTCGCCAGCGTCCCCAAGCACGTGCCCTGCACCCCAGA
15 GAGGCGTCCCCGCACTGGGGCTGGCGGGGAGGGTGCGGGGAGTGGTCCCCCTGTCTCGCTTCCAGCCCAG
AACCATCTCTTCTCTCCCATCCCTGCCCTCGGCCCCACAGTCCCACGCTGCTGGCTCCAAGCAGCACGAG
AGCATCCCGGGCAAGGCCAAGAAGCCCAAAGTGAAGAAGAAGGAGAAGGGCAAGAAGGAGAAGGGCAAGA
AGAAGGAGGCTCCCCACTGAAGGGCCCTGGACAGGGCTCATTAAACCTTCCTCTCTGCCTTC

20 >DEX0367_044.nt.1
AAAATTAGGCTAGGTACGAGGCYTGGGTGTAATCGGCCGAGGCGAGCGACCTCCGCGCGTTGGGAGGTGT
AGCGCGGCTCTGAACGCGCTGAGGGCCGTTGAGTGTCGCAGGCGGCGAGGGCGCGAGTGAGGAGCAGACC
CAGGCATCGCGCGCCGAGAAGGCCGGGCGTCCCCACACTGAAGGTCCGGAAAGGCGACTTCCGGGGGCTT
TGGCACCTGGCGGACCCTCCCGGAGCGTCGGCACCTGAACGCGAGGCGCTCCATTGCGCGTGCGCGTTGA
25 GGGGCTTCCCGCACCTGATCGCGAGACCCCAACGGCTGGTGGCGTCGCCTGCGCGTCTCGGCTGAGCTGG
CCATGGCGCAGCTGTGCGGGCTGAGGCGGGAGCCGGGCGTTTCTCGCCCTGCTGGGATCGCTGCTCCTCTC
TGGGGTCCTGGCCGGCCGACCGAGAACGCAGCATCCACGACTTCTGCCTGGTGTCGAAGGTGGTGGGCAGA
TGCCGGGGCCTCCATGCCTAGGTGGTGGTACAATGTCACTGACGGATCCTGCCAGCTGTTTGTGTATGGGG
GCTGTGACGGAAACAGCAATAATTACCTGACCAAGGAGGAGTGCCTCAAGAAATGTGCCACTGTCACAGA
30 GAATGCCACGGGTGACCTGGCCACCAGCAGGAATGCAGCGGATTCCTCTGTCCCAAGTGCTCCCAGAAGG
CAGGATTCTGAAGACCACTCCAGCGATATGTTCAACTATGAAGAATACTGCACCGCCAACGCAGTCACTG
GGCCTTGCCGTGCATCCTTCCCACGCTGGTACTTTGACGTGGAGAGGAACTCCTGCAATAACTTCATCTA
TGGAGGCTGCCGGGGCAATAAGAACAGCTACCGCTCTGAGGAGGCCTGCATGCTCCGCTGCTTCCAGGGG
AAYYTCCCGGCSCTKGKCCARGGGGGGGCCGGGGGGCGGCCCGCGGGGGGACCCCTCGGGGCGCCCAGGGG
35 ACAGAACCGGG

>DEX0367_044.nt.2
CCACCGCTGGGGTCCCCAGCATTAAGCCCTGGAGCATGAGCCCAGTGAACTTTGAAATTCAGGAAGGAATT
CACAGGAAAGCTGCGGCGAGGCGGGCAGAAAACCACAGGCTTCCAAGAAACGCGTGTCTGCCGCCCCCTC
40 CTCCCACCCTGAGGCCAGTCCTAGAGCGAAGGCCCTCCTCCGACGTCGTCACCAAGCCCAAGGGAAGGGT
GGCAGGTGCTCAGCGGGCAGACGCCCCGCCCCGCCCCGCCAGGTTCTGTTGGGGGCGAGGCCCGCGCAAG
CCCCGCCTCTTCCCCGGCACCAGGGGCGGGCCCAGGTGCGCCCAGGGCCGGGGAGCGGCCGCGCAGGTGC
CTGCCCTTTGCGCCTGCGCCCAGCTCGCCCTGCCTAGCCAGGTGCGCCCCGCCCCCTGCCTGCCCGGCCA
CCTTCGGGAGCCGCTTCCAATAGGCGTTCGCCATTGGCTCTGGCGACCTCCGCGCGTTGGGAGGTGTAGC
45 GCGGCTCTGAACGCGCTGAGGGCCGTTGAGTGTCGCAGGCGGCGAGGGCGCGAGTGAGGAGCAGACCCAG
GCATCGCGCGCCGAGAAGGCCGGGCGTCCCCACACTGAAGGTCCGGAAAGGCGACTTCCGGGGGCTTTGG
CACCTGGCGGACCCTCCCGGAGCGTCGGCACCTGAACGCGAGGCGCTCCATTGCGCGTGCGCGTTGAGGG
GCTTCCCGCACCTGATCGCGAGACCCCAACGGCTGGTGGCGTCGCCTGCGCGTCTCGGCTGAGCTGGCCA
50 TGGCGCAGCTGTGCGGGCTGAGGCGGGAGCCGGGCGTTTCTCGCCCTGCTGGGATCGCTGCTCCTCTCTGG
GGTCCTGGCCGGCCGACCGAGAACGCAGCATCCACGACTTCTGCCTGGTGTCGAAGGTGGTGGGCAGATGC
CGGGGCCTCCATGCCTAGGTGGTGGTACAATGTCACTGACGGATCCTGCCAGCTGTTTGTGTATGGGGGCT
GTGACGGAAACAGCAATAATTACCTGACCAAGGAGGAGTGCCTCAAGAAATGTGCCACTGTCACAGAGAA
TGCCACGGGTGACCTGGCCACCAGCAGGAATGCAGCGGATTCCTCTGTCCCAAGTGCTCCCAGAAGGCAG
55 GATTCTGAAGACCACTCCAGCGATATGTTCAACTATGAAGAATACTGCACCGCCAACGCAGTCACTGGGC
CTTGCCGTGCATCCTTCCCACGCTGGTACTTTGACGTGGAGAGGAACTCCTGCAATAACTTCATCTATGG
AGGCTGCCGGGGCAATAAGAACAGCTACCGCTCTGAGGAGGCCTGCATGCTCCGCTGCTTCCGCCAGCAG
GAGAATCCTCCCCTGCCCCTTGGCTCAAAGGGTAAGTGGCCCCTTACCCTCCTCCTGCCATCAGCCTGCC
TCCTCCCTTCCTTGACTGAGCTCAGCCCTGCCCAGCTGTGGTTTACATTATCCTTCACTGTGAACATCAT
CTTGGCAGAAAGTCATGTTTCTGCGTGAGAATGGCGAGGTGGTGGTTTGTCCCACCGTTCAGTGTACACA
60 GTTGGGGCTGGAGTGAGTCAGTCACAAGGCAGGCCCTGCCCAGGCGGCGTGGGTGACTGGGGATGAGGTC

TTCCTGTTGAGCATTTGAGGACTGCTGCACACGGGCCTGAGGCTGGCCTGAGGTGTGGAGGGAGC

>DEX0367_045.nt.1
AGTGTGATGGATCGGCCGCCCGGGCAGGTACCTGGCCATCTTGGGCAGTGTGACGTTTCTGGCTGGCAAT
CGGATGCTGGCCCAGCAGGCAGTCAAGAGAACAGCACATTAGTTCCAGAAGAAAGATGGAAATTCTGAAA
ACTGAATGTCAAGAAAAGGAGTCAAGAACAATTCACAGTATGAGAAGAAAAATGGAAAAAAAAAACTTTA
TTTAAAAAAGAAAAAAGTCCAGATTGTAGTTATACTTTTGCTTGTTTTTCAGTTTCCCCAACACACAGCA
GATACCTGGTGAGCTCAGATAGTCTCTTTCTCTGACACTGTGTAAGAAGCTGTGAATATTCCTAACTTAC
CCAGATGTTGCTTTTGAAAAGTTGAAATGTGTAATTGTTTTGGAATAAAGAGGGTAACAATAGGAAAAAA
AAAAACAAAAAA

>DEX0367_045.nt.2
TGATGGATAGGTCGCGGCCGAGGTACCTGGCCATCTTGGGCAGTGTGACGTTTCTGGCTGGCAATCGGAT
GCTGGCCCAGCAGGCAGTCAAGAGAACAGCACATTAGTTCCAGAAGAAAGATGGGAAATTCTGAAAACTG
AATGTCAAGAAAAGGAGTCAAGAACAATTCACAGTATGAGAAGAAAAATGGAAAAAAAAAACTTTATTTA
AAAAAGAAAAAAGTCCAGATTGTAGTTATACTTTTGCTTGTTTTTCAGTTTCCCCAACACACAGCAGATA
CCTGGTGAGCTCAGATAGTCTCTTTCTCTGACACTGTGTAAGAAGCTGTGAATATTCCTAACTTACCCAG
ATGTTGCTTTTGAAAAGTTGAAATGTGTAATTGTTTTGGAATAAAGAGGGTAACAATAGGAAAAAAAAAA
ACAAAAAA

>DEX0367_046.nt.1
AAAATNTAAGAAGGTTCGCAGACCTGACATCCAGTACCCTGATGCAACAGACGAGGACATCACCTNACAC
ATGGAAAGCGAGGAGTTGAATGGTGNATANAAGGCCATCCCCGTTGCNCNAGGACCTGAACGCGCCTTCT
GATTGGGACAGCCGTGGGAAGGACAGTTATGAAACGAGANAGCTGGATGANAAGAGTGCTGAAANNAAAA
NNAAAAAGMAGTCCAGATTATATAAGCGGAAAGCTAATGATGAGAGCAAWGAGCAKRCCGATGTGATTGA
TAGWMAGGAACTKKCMAAAGRMAGCCGTGAAKYCMACAGCCATGAATKKNACAGCMAKGAAGATATGCTG
GTKGTAGACGCCAAAAGTAAGGAAGAAGAWAAACACCTGAAATTTCGTATATCTCATGAATTAGATAGTG
CATCTTCTGAGKTCAATTAAAAGGAGTAATAAAATACAATTTCTCACTTTGCATTTAGTCAAAAGAAAAA
ATGCTTTATAGCAAAATGAAAGAGAACATGAAATGCTTCTTTCTCAGTTTATTGGTTGAATGTGTATCTA
TTTGAGTCTGGAAATAACTGATGTGTTTGATAATTAGTTTAGTTTGTGGCTTCATGGAAACTCCCTGTAA
ACTAAAAGCTTCAGGGTTATGTCTATGTTCATTCATAGAAGAAATGCAAACTATCACTGTATTTTAATA
TTTGTTATTCTCTCATGAATAGAAATTTATGTAGACGCAAACAAAATACTTTTACCCACTTMAAAAGAGR
ATATAACATTTTATGWCACTATAATCTTTTKTTTTTTAAKTTAGTGTATATTTTATTGTGATTATCTTTT
TGTNTG

>DEX0367_047.nt.1
AAGAGTGGTATCAACGCAGAGKGCCCATTACGGCCGGGGGCCGCCATCTTGGGGCTGCTGGGACTCGCGT
CGGTTGGCGACTCCCGGACGTAGGTAGTTTGTTGGGCCGGGTTCTGAGGCCTTGYTTCTCTTTACTTTTC
CACTCTAGGCCACGATGCCGCAGTACCAGACCTGGGAGGAGTTCAGCCGCGCTGCCGAGAAGCTTTACCT
CGCTGACCCTATGAAGGCACGTGTGGTTCTCAAATATAGGCATTCTGATGGGAACTTGTGTGTGTTAAAG
TAACAGATGATTTAGTTGTGTTGGTGTATAAAACAGACCAAGCTCAAGATGTAAAGAAGATTGAGAAATT
CCACAGTCAACTAATGCGACTTATGGTAGCCAAAGGAAGCCCGCAATGTTACCATGGAAAACTGAGTGAA
TGGTTTGAAATGAAGACTTTGTCGTGTACTTAGGGAAGTAAATATCTTTTGAATTAGAGAAAGTGTTGGG
ACAGAAAGTACTTTATGTAACTAAGTGGGCTGTTCAGAAGCTTAGAGGTCATTTTTGTAATTTTCTTTTT
AAATTACTTTAGAGAGCTAGGGATGCAAATGTTTTCAGTTAGAAAGCCTTTATTTACTTTGTGGAAATTG
AACAAGAAATGCATCTGTCTTAGAAACTGGAGATTATTGATGTTAGGTAACACATGTAATGGTTCTCGTG
GCAATTGTGTCTCAGTATTTGGAAATGAGATTTACGGAAAACCATTTTTTAAATTAGTCATGCTTCTTTA
ATGACCCTAGTGTACACTTGGTGATCAGGTTTGGGA

>DEX0367_047.nt.2
AAGAGTGGTATCAACGCAGAGKGCCCATTACGGCCGGGGGCCGCCATCTTGGGGCTGCTGGGACTCGCGT
CGGTTGGCGACTCCCGGACGTAGGTAGTTTGTTGGGCCGGGTTCTGAGGCCTTGYTTCTCTTTACTTTTC
CACTCTAGGCCACGATGCCGCAGTACCAGACCTGGGAGGAGTTCAGCCGCGCTGCCGAGAAGCTTTACCT
CGCTGACCCTATGAAGGCACGTGTGGTTCTCAAATATAGGCATTCTGATGGGAACTTGTGTGTGTTAAAG
TAACAGATGATTTAGTTAGACAGTGTCTTGCTCTATGTGCTCCAGGCTGCAGTGCAGTGGCTTGATCATG
GCTCACTGCAATCCTCGACCTCCTGGGCTCAAGCGGTCCTCTTGCTTCAGCCTCCTGAATAGCAGAGACT
RCAGCGGGTAGCTGGGATTACAGGCACCTGCCACCACACCCAGCTAATTTTTATATTCCTAGTAGAGACA
GGGTTTCACCATGTTGGCCAGGCTGGTTTCGAACTGCTGATCTGGTGGTCCGCCTGCCTCGGCCTCCCAG
AGTGCTGGGATTACAGGCGTAAGCCACCACGCCTGGCCAAGAAAATCAAGATTTACCTGGWTTATGTGTT
GACCAGCTATACACAGAGAATTCTAGATTTTKTTCTAAAAATTATTATCAAACCCSSAATTTCACCAGTA

GAAAAAGAGATCTTAAGATTTCTGTGTGTTTTTTTTTCCAGCATAACTCAGTTACTTATGGCTGGGAGAAAA
TATGTAGGGAGATAATCTAATACCTAATTCGAAGGRAGAAAATATCTAATTCTAAGTCTTGAGTTCTTCT
YATYMGACCATAGCAGGSTTTAACATTTWACATTTTTTCARAAGCTTCTAAWAGACYTCATTCARTTTTTT
GCCMCAGCAGKGGCATTTAAATMACTGRCTTTAMARCCATGAATYGCTGGAGGSTTGGAAACTTGAGTCA
5   ATTTCCCTAGTACSACCTCCAAGGCAGAACTAGCAGCAAAACAGTGGTAGGAAAAACTCTCAAAATTTTA
CCCAATCTGTATGTTTCTCTACATTGTCAGTATCTAGTTCTTATATAGTCTCACATCATGTCACTTCTAA
CAATTTCCTGACAGTGTTGGTGTATAAAACAGACCAAGCTCAAGATGTAAAGAAGATTGAGAAATTCCAC
AGTCAACTAATGCGACTTATGGTAGCCAAAGGAAGCCCGCAATGTTACCATGGAAAACTGAGTGAATGGT
TTGAAATGAAGACTTTGTCGTGTACTTAGGGAAGTAAATATCTTTTGAATTAGAGAAAGTGTTGGGACAG
10   AAAGTACTTTATGTAACTAATCTTCTGATCATCCCACTTAGTTACATAAAGTACTTTCTGTAACTAAGTG
GGCTGTTCAGAAGCTTAGAGGTCATTTTTGTAATTTTCTTTTTAAATTACTTTAGAGAGCTAGGGATGCA
AATGTTTTCAGTTAGAAAGCCTTTATTTACTTTGTGGAAATTGAACAAGAAATGCATCTGTCTTAGAAAC
TGGAGATTATTGATGTTAGGTAACACATGTAATTGTTTCTCTGGCAAATTWGTATCAGTAATTTGAAAAT
GAGATATTAGGAAAAACAATTCTTCTTAAATTTAGTTCATCTTTCTTTCAACACAGCAACATATCAAATG
15   TAACCCATATTAGTCCAGATCCATGTATTATGGAGCATACAAATGTATGCTGTAGTGACCAATAAATCAT
AACATATGGTAATTGGCACTTAACTCCACACCACTAGTATGCACTTGTTCATACACTACTGTGTACCTAA
TTATAGTATACGCAGTGTAGTCTCAATTAATAATCTGAAACGTAACTTAGTGGCCTAAACAAGTATGCAA
TAGAACTTATTCTACCACATATCAAACACTACCCATGCAAAACTACTACACGGGGGATACACGCAACACA
ACTACAAGACAATAGACCAGAACGCAACGAGGCACTATAGGACAAAACATAAGTCACACACACCAACGAC
20   ACACGAAACCCCCCACACACGCAATAACTAGACCACTGATACGGACACTGGCGGAACCTGACTGGAAATC
CACAAGATCGACCAAATACATCGGGAACCCGCCACACCGAAACCCAAAATTAGACAAGGCAAAAATCACA
GCAACGACACCAATAAATAAGAAGCCCAACACTAGACACACAAGAAACCCCCCACCAAACCGGCGACCAA
TGCCAATACACCAGCAAACGCACACCAAGGACAAAAAAAGACACGAATGACACAAAACCGTCCCGAGGGA
AAAAAACACAAACAGAACACAAAAAAAAACACAAGAAACACACAACACTAAAAGCCCACAAAAACGCAACA
25   AAACAAAGAAAACAAAGCAAAAACAAAAACGGCAACCCCACTGGCAACACAAACAGAACAAAAATGAGAC
AAAGTTCACGAAGAGCTAAGAGAGAGAAAAGCAAAGCACAGGGAACCATGGAACAGACCCACATACAGTG
CAAACAGGGACGGCAACAAACACAAAAAGAACAAGAACGAGCACCGAGAAGCACACCCTAAAGAAAAGCA
AAAAGAAAACAGAACA

30   >DEX0367_047.nt.3
AAGAGTGGTATCAACGCAGAGKGCCCATTACGGCCGGGGGCCGCCATCTTGGGGCTGCTGGGACTCGCGT
CGGTTGGCGACTCCCGGACGTAGGTAGTTTGTTGGGCCGGGTTCTGAGGCCTTGYTTCTCTTTACTTTTC
CACTCTAGGCCACGATGCCGCAGTACCAGACCTGGGAGGAGTTCAGCCGCGCTGCCGAGAAGCTTTACCT
CGCTGACCCTATGAAGGCACGTGTGGTTCTCAAATATAGGCATTCTGATGGGAACTTGTGTGTGTTAAAG
35   TAACAGATGATTTAGTTAGACAGTGTCTTGCTCTATGTGCTCCAGGCTGCAGTGGCTTGATCATG
GCTCACTGCAATCCTCGACCTCCTGGGCTCAAGCGGTCCTCTTGCTTCAGCCTCCTGAATAGCAGAGACT
RCAGGTGTTGGTGTATAAAACAGACCAAGCTCAAGATGTAAAGAAGATTGAGAAATTCCACAGTCAACTA
ATGCGACTTATGGTAGCCAAAGGAAGCCCGCAATGTTACCATGGAAAACTGAGTGAATGGTTTGAAATGA
AGACTTTGTCGTGTACTTAGGGAAGTAAATATCTTTTGAATTAGAGAAAGTGTTGGGACAGAAAGTACTT
40   TATGTAACTAAGTGGGCTGTTCAGAAGCTTAGAGGTCATTTTTGTAATTTTCTTTTTAAATTACTTTAGA
GAGCTAGGGATGCAAATGTTTTCAGTTAGAAAGCCTTTATTTACTTTGTGGAAATTGAACAAGAAATGCA
TCTGTCTTAGAAACTGGAGATTATTGATGTTAGGTAACACATGTAATTGTTTCTCTGGCAAATTWGTATC
AGTAATTTGAAAATGAGATATTAGGAAAAACAATTCTTCTTAAATTTAGTTCATCTTTCTTTCAACACAG
CAACATATCAAATGTAACCCATATTAGTCCAGATCCATGTATTATGGAGCATACAAATGTATGCTGTAGT
45   GACCAATAAATCATAACATATGGTAATTGGCACTTAACTCCACACCACTAGTATGCACTTGTTCATACAC
TACTGTGTACCTAATTATAGTATACGCAGTGTAGTCTCAATTAATAATCTGAAACGTAACTTAGTGGCCT
AAACAAGTATGCAATAGAACTTATTCTACCACATATCAAACACTACCCATGCAAAACTACTACACGGGGG
ATACACGCAACACAACTACAAGACAATAGACCAGAACGCAACGAGGCACTATAGGACAAAACATAAGTCA
CACACACCAACGACACACGAAACCCCCCACACACGCAATAACTAGACCACTGATACGGACACTGGCGGAA
50   CCTGACTGGAAATCCACAAGATCGACCAAATACATCGGGAACCCGCCACACCGAAACCCAAAATTAGACA
AGGCAAAAATCACAGCAACGACACCAATAAATAAGAAGCCCAACACTAGACACACAAGAAACCCCCCACC
AAACCGGCGACCAATGCCAATACACCAGCAAACGCACACCAAGGACAAAAAAAGACACGAATGACACAAA
ACCGTCCCGAGGGAAAAAAACACAAACAGAACACAAAAAAAACACAAGAAACACACAACACTAAAAGCCC
ACAAAAACGCAACAAAACAAAGAAAACAAAGCAAAAACAAAAACGGCAACCCCACTGGCAACACAAACAG
55   AACAAAAATGAGACAAAGTTCACGAAGAGCTAAGAGAGAGAAAAGCAAAGCACAGGGAACCATGGAACAG
ACCCACATACAGTGCAAACAGGGACGGCAACAAACACAAAAAGAACAAGAACGAGCACCGAGAAGCACAC
CCTAAAGAAAAGCAAAAAGAAAACAGAACA

>DEX0367_048.nt.1
60   GCAGGGCTTCCACTACACCCCCACGTGGCGAGCGCCATCCAAGACCTTGTGTTCAGGCAGAGGCTGGGTC

GGGCCCCCACCTGCCTGGGCCTGGCCCTCTAAGCCCCCTTGGGCTCCCCTGCTGCTCTCTGTGGTGAAGC
CTGTGTTGCTGGTGGAGCTTTAGCCAGGTGAGTGCAGCCAGGTTAAGTTTCTCCAGGTGTCCAAGGGGAC
AGCTGCACGGAGTAAAAAGGCAAGTGGAAGGATGTGGGCCAGAAGGGCAAGCACAATAAAGCTGAAATTC
ACCACTAGCAAAACATCCTCAGAGTCACAAGAGCCAAGCCTGAATCAGCGTTGCTCTTTCACAGCCCAAG

5   GGTGACCCGAAAACGCCCATCCTGGGCAGCCTTCCTCGACTGGCAGGGACTCAAGGGCCGGAAGTGCGAG
GTGGCCGTGCCAGGCTGTGCCAGCATGCACCAGGGCAGGGAGCATTCTGCCCAAACTATTGACTCGGACT
ACCTGCTCACAGTCTGCAGAAGGCACGCTGGCGGCCTCAAGCGGGGCTGCATCCTGCTTCCTTCTCTTGC
TGGGTATGGCATTAGCAGCTTCCCATTCATCGGATATGAGGTCTACAGCAGCATCTACCAAAACCTCTGA
TGTAGCTCCCTGTTGAAACACGACCCCCTTCAGACAGGGCAGGGCTGTGCACCCCACWGCACAGCTCCAG

10   TAGACTCGCAAAGACCCCACCTCAGGGGCCACCTCCTCTTGGGAAGGCCCAGGCAGGGACACCCACATGT
GGGCACTGACCAGCTTCCCAGCACCGTCTGCCTCCAGGGCCCATGGGCAGGAGGGTGTGGGAAGACGAAG
AAGCAGCTCCTGCAGMCGTGCAGGACAAGTCCACGTAAGAGGAGAACTGCCCACACCAGAGGCGGCAGCA
CTGCCAGGCTCACG

15   >DEX0367_049.nt.1
AGCGGAGGGTTCACCCTCCGCTCCACAGGGTCGGCAGCAGGGCGGGGCCTCCGGAAGCTCCGCCCCACGC
TTTCCCGGGGCGCATGCGACGTGGGCCGGAGCGTCTGGAAGCTCCGCCCGTCGCACTGTAGAGTCGGCCG
AGGCGCACGAGGTATTTTTCACGCTCCGCCCCTCTGCAGGCTAAAGTGCGTGGGCGGGAAGCGGTGGGCA
GGGTGCCATCTGGCTCCGCCCTTCTCCTGTGGTGTGGGCCAGGCGGCGGCTTCCTCCTCCTGCAGCAGCC

20   ACAGGCTCCACTCTGATCCTTCTCCCGCGGCATGGATCCTTCTCCCGCGGCGTGGATCCTTCTCCCGCAA
TCTCCGTGCGCGTCCCCAGTCAGTACCCGCAGCCTCCCGACGCACCCGCTGGCTCCAAGCCTCCCTACCC
CAGGTTTCCTGGCTAAGAGAGAGACAGAGGGAGAGAGGGGGGAAGACGAGAGAACAGGCAATGGGAGGTG
ATGGTGAGAGCTTATTGGAACAGACAAGAGGGAGGAAACCCACATTCCTTCATTCCCCATCCATTCATTT
ATGGCCTTATGTATTCCATTGAATCTTCACAGCTCTAGAAAAAAGTGTGCTACAATTATTCCCTTTATTA

25   CATGAGGTCACTGAGGCACAGTTTAAGAAAAATTTGCCAGCAGGCACAGTGGGTCACTCCGGTAATTCGCA
GAACTTTGAGAAGGGGGAGGAAGGTGGATCCATTGAAGCCCAGGAAGTTGGAGACCAGGTCTGGGCAACA
ATGGGGGAGACCCCGTTTCTACAAAAACTTAGGCAAAATTAGCAAACTGGTTCACACTTGTAGTCCCAGG
GTATTCA

30   >DEX0367_050.nt.1
CTTCGGGCCTAACCACAAGAGGCATGTCTGGTTCCGCGAGAGCATGACCGATTGGATTCCAGTTCGAGTA
TGCCGCCTAGGGCTCCGATCCTGCCGATGTGGCCTATCCAGCTGACCTTTCTGCGCCTGATGTCCACTGG
AGGCCTCCCAGGACATCACCTACCACTGCAAGAACAGCGTGGCCTACATGGACCAGCAGACTGGCAACCT
CAAGAAGGCCTGCTCCTCCAGGGCTCAACGAGATCGAGATCCGCGCCGAGGGCAACAGCCGCTTCACCTA

35   CAGCGTCACTGTCGATGGCTGCACGAGTCACACCGGAGCCTGGGGCAAGACAGTGATTGAATACAAAACC
ACCAAGACCTCCCGCCTGCCCATCATCGATGTGGCCCCCTTGGACGTTGGTGCCCCAGACCAGGAATTCG
GCTTCGACGTTGGCCCTGTCTGCTTCCTGTAAACTCCCTCCATCCCAACCTGGCTCCCTCCCACCCAACC
AACTTTCCCCCCAACCCGGAAACAGACAAGCAACCCAAACTGAACCCCCTCAAAAGCCAAAAAATGGGAG
ACAATTTCACATGGACTTTGGAAAATATTTTTTTTCCTTTGCATTCATCTCTCAAACTTAGTTTTTATCTT

40   TGACCAACCGAACATGACCAAAAACCAAAAGTGCATTCAACCTTACCAAAAAAAAAAWWWAMAMAARGA
TTAWTTAATTTTTTTTTSMAMAAGGAAGCTGGGCCCCCTGGCTTGAAAACCCATGCGGGGGTAGGTCCCTT
TTGGCCCGTGGGGTTTTTGAAACCCCAATGGCGTGCCCTTTCGAGCTCCTTTTCTCAAAACCCCCCTGGG
GGCCTTCCCCTCCACTCCTTCCCAAATTCTGGTCTCCCCAGAAGACACAGGGAAACATTGTTTGGTCTGC
CCAGCAAATCAAGGGCAATGCTCAAACACCCATGGGGGCCCCAACCCTTCAGCCCGTCCGTGCCCGCCAA

45   AGAACCCCAAGGCCCTGGGGGACATGGGGCTCGCTTCAGAGTTGCCAAAAAAAGCCTTGTGCCAAATCAG
GGGGCCTCCCAAGGGGCTCTGTTGAAAACATTGTCCCCTACGGTGATATGGAGGGGGGTCAATGCCGGGG
GGAAGCCACCAAGGCCCCATCAACGGGGATTCCAGAAGGAGCAGCCAGCAAGGGGCCCACGGCTACGAAA
TAACCGCCGAGAAACCACCGGTTGGGCGGAACCGCGTGTCAAATACCCTTGGTTGCCCGCAAGGGACGGG
GACCGGCGCAGAGCCATGGTCGTGTACTTGGTGTTACTGGTGTGCGCGGAAAAAATACCCCGGCCTGGTCC

50   TGGAGGTGCCCCGCGACAAATTGACTGCGGCGGGAATGAAGCACAGATGAAGGACCCCAT

>DEX0367_051.nt.1
GGTGATTGCCAGTCCTCCGGAGTGGAGATCCCAGKTCATTTAAAGTGGTCTAAGTKCCGGATGACATAGG
GGTAGATTTCCTTATGGTTCCAGAGCCCAAAATCATTGATGCTGCTTTGCGGGCATGCAGACGGTTAAAT

55   GATTTTGCTAGTACAGTTCGTATCCTAGAGGTTGTTAAGGACAAAGCAGGACCTCATAAGGAAATCTACC
CCTATGTCATCCAGGAACTTAGACCAACTTTAAATGAACTGGGAATCTCCACTCCGGAGGAACTGGGCCT
TGACAAAGTGTAAACCGCATGGATGGGCTTCCCCAAGGATTTATTGACATTGCTACTTGAGTGTGAACAG
TTACCTGGAAATACTGATGATAACATATTACCTTATTTGAACAAGTTTTCCTTTATTGAGTACCAAGCCA
TGTAATGGTAACTTGGACTTTAATAAAAGGGAAATGAGTTTGAACTGAAA

60

>DEX0367_052.nt.1
CTGTAGTGATCTGAGGATCCCTTCGATTGGTGTGGGAACTTATTGGTGCTCTGAGTAGAGGGCCTGATGA
GGTTGACAGTAATGATTGAAAGAGGCATTTGAGTGATTCACAGTACTAATGCGGGTTCTCGGCCCTATGT
TGCCACCAGCAACTGCACAGCTCTGTCTGTTTGACCCAAGAGCAGTAGCTGGTCTGTACATTTCCTTGTA
TGATCATAAGTTGCCCGCCAATCGAATCGATTTGCGTGGACAACTTAACTTATAACATTTTAAGTAAAAT
CATGCATTTTAGGCAGACAGACATCTCACTTTTTAGTATATATTCCTTCTTCTAGAGGATTATTTGGTGT
GTGTCTGCGTGTGACCGTGTGTGTTATTGAAAGGAAACAGTTCCGTGTTTAGAGTTTTGGGTTGTTAGT
TAGAGGTCTAGTTAAGCTATAAAATTACAGTGGAATAAGTGATGACAGCCAGTTAGTACAGTTCCCTCTT
ATAGCTGAACTGAGCTTAACACACCAACAGAAATGTTTGGCAAATAGACTGCTAAAATGACTCTTTCCTT
TGTGGGGACATAAGTTGCCTGACATTTACATGTTTATAATGAGTCCTGGAAGTAGAAATGAAAATATGTG
CAGAAAATAGATTTCTATCATCACTTCCTTTTCTTATTGAGGGTAGCATTGTCCAGTACGTTTTTGTCAT
CTCTGTGCTTTCTTGTTGTCCACCCATACAGGAAGAATGTTGGTGAAAAGCACAGAAACCAGTGACGTTT
GTGGCTGTGGAATTTCCATGGAGAAAAGAGAGCATCTGAACACCTGGACCCATCTTTTGCACCTGGCAGAC
CCTCTGCACTCACCCCAGCGTGTTCTGTGAACTTGAGTGACAACGCGTGCTTGCAGGGTGCTTTTTGGAT
GACTGGGGAAGAGGTGGGGAGGGGGTGGTGGGGGGGAAGCATGGACGAGAACATGGAGCAAATGTTTTACA
ACCTGAACCTCAGAACTGTGATCCTCCAAGGAGCGCGCTACTTGAAGAAAAGAAAAAAAAAAAAAAAAGT
CGAATGGCTTCTCAGGGATTTTGTTTTCCGTGCACATATAAGCCATAGTTACAGTACAGGTGGCAGTATT
TAGAGCACTCAGATTTCAGTTCTGTTAAATGTGAAAGAGGGATCGACTGACGTTCATTGCTGTTCCATAA
CTAGTGTAAAATATGTATATGTTTATCTTTATTTTTATAATATGCAAATACATTTAAATTTATACAGTTT
GAAGCTTCTAGCGTTAGTTCACACTGGGTGCAATATTCTGCATGAGAACTGTGCCAAGATGGGCTGATT
TCTTATTTTAGTATAAGACTTTTTTGTTTTTTCTTTATTCTTTAATCCTCTTCACATTAAAAAAAAAAAA
TCTCTCTGTTAGCCCATGGATTAAGTGTTGGTTCATAGAGATTGCCAATAATCAGAAAGAACCTTAAATG
TGCATTTAAGACAGTGTCCCTTCCCTTCTTTTCAATGAAGGTCCCTGCCTATATAAATCATCTGGCACGC
TGGTGGGAAATCCTTTGCTCTTCCAACGTGTTATTAGTGCTGGGCAGAGATGGGGCACACTCAGGGGCCA
AAGAGGACAAAAAGTCCATGCAAAACTTGAGTCTTTTAATGGCTTAAGATAATCAGGAGTCAGTTCTGAA
TCTTACAAAGTGCTCTGCTTAATAAGTACCTTACTTAGCAGAGCACTTTGCAAACATATTACTTATTAGC
AGAGCTCTTTGTAGACCTTCCACATCTGGCTGTCAGATCTTAAGGTTGTGAATTTAGGCTCCAGTTATAT
TCACTGGAGAGCATAATCCCACACGGGTTATTTATAAATACAGAGCCTCTGATTGGACGGTCTCCTGCCA
AGAACTAGTAATACCCTTGTTTTAAAATCTTCACAAGGTAAAACTTAAAAAGCCAACCAAACAAATTGCT
CTCCATTCTACTTTTAATTGGGCCAAACAGCATATGCTACAGTAGTAACATGTTTTTCGGAGAGTGTAAA
AAACTCTGTTTACATTTGCCTCCTCCGTGGGTTGATCGAAAATGTATAAAACTGACTGCTTCTCGCCAGC
CTCAGACAAGAAGAGTGAGTTGCTGGTACTCGCTACTCTTTTACTTCTTTTGTAAAGTATTGACTCTTGG
AAGGCTACAGTATACAAAGTCTCAACATGTTTTTTAAAAGAAATAAGGAGCAAGCGACTGCCCTGCTAGA
AATCACAAACCGATTTTTGTAGAATATTTTGTGCCCCAGGCATTAATTTCACTGACTCCAGAACCTGCAG
TTCAGAGAATGATTTCTTATGATGATAAAAATCGAATGGGATCAGACGATGTTTGCATTTTTTTAATACT
TGAATAGGACACCTCAAGTTTGAGATTTCATTTTTCTTTTAGAACACAGTCACAAGATTAATCTGGTGAAT
CCTTTTGTCACAGTTCTCGTGTGTGTGTGCGCGTCTCCGTGTGTGTGTGTGTGCATGTGTGTAAAACTGA
ATGGTCACATTTAATTGCTTTTTGGACCATTGAATAGTTGGGAAGTAAGAATTTTTTAATTGGCATGAGA
CGGTTCCTCAACTGTTAAAATTAACCAACTTTGACCTGTCTTTAGAAAAAGGCTTATTTGTATGATTTTGG
GCTAACTCCCCGGGGACCATATTAAATGACAAAAATGCTCCTTTGGGTGACACACCCTACAAAGTATTTG
CTGTTACGAACATAAACGCCCACATTCTTAATATCTAATATTTTTGACCAGTGATGTTTTATGCTGTCAT
CTGAACCCTAGAGAAGCAGTGTCAGAGGAAACCTTGGTGTCACATGTGTCTTAGCAAAAGGGTTACCATG
ATCGAGGGTCATGTGACCAAAGATGCTCCAGAGAAGCTTGAGAATTTGTTTCAAGTTGGGAGGAGGGTT
GGAGATACAAAAATCACTCTGCTCTACAGGACTCTTCAGCTGTCTATGCAAGAAATTCCGTTTTCTCTTT
CAGCACCTGGAAAGACACAGCAGCCCACCGAGGCGATAGGTGATTCACTAAGCACAAGAGGAATGTTTTC
TAAGCAAGGCGTCCCTTGCCTCTCAAACAAATGCCCTCCAAGTTTGTTAGGGGTTTCTATTCCTGCAACTT
GTGGTATCAAAACCACTTCCTGGAATTGTCAAAGCACTGCCAAAATAAATGTTTTTCCCCCTTCTAAGAA
AAAAAAATGACAGTGCTCATATTTGACACTTGTGTATTGGACTCTCTTTTGAATGAATAAAAAGGAAAA
GGGGTTTGGTGTAATTCCTGATGGGGTGCAGAAATTAAAAAAGAAAAATCAATAAAAATACAAAGAAATGG
TTTCCCATTTCGTTGTTGTAACTGGGCAGAAATTAAAAAAGAAAAATCAATAAAAATACAAAGAAATGG

>DEX0367_052.nt.2
ATCAATGAACCCGCAATCATGAAACTTCTGACACCACCAGGAGCTGACTCACACCTGAACAGTACACCGT
CGAAGCTTCGACGGTCTCCTGCCAAGAACTAGTAATACCCTTGTTTTAAAATCTTCACAAGGTAAAACTT
AAAAAGCCAACCAAACAAATTGCTCTCCATTCTACTTTTAATTGGGCCAAACAGCATATGCTACAGTAGT
AACATGTTTTTCGGAGAGTGTAAAAAACTCTGTTTACATTTGCCTCCTCCGTGGGTTGATCGAAAATGTA
TAAAACTGACTGCTTCTCGCCAGCCTCAGACAAGAAGAGTGAGTTGCTGGTACTCGCTACTCTTTTACTT
CTTTTGTAAAGTATTGACTCTTGGAAGGCTACAGTATACAAAGTCTCAACATGTTTTTTAAAAGAAATAA
GGAGCAAGCGACTGCCCTGCTAGAAATCACAAACCGATTTTTGTAGAATATTTTGTGCCCCAGGCATTAA
TTTCACTGACTCCAGAACCTGCAGTTCAGAGAATGATTTCTTATGATGATAAAAATCGAATGGGATCAGA

```
CGATGTTTGCATTTTTTTAATACTTGAATAGGACACCTCAAGTTTGAGATTTCATTTTCTTTTAGAACAC
AGTCACAAGATTAATCTGGTGAATCCTTTTGTCACAGTTCTCGTGTGTGTGTGTGCGCGTCTCCGTGTGTGT
GTGTGTGCATGTGTGTAAAACTGAATGGTCACATTTAATTGCTTTTTGGACCATTGAATAGTTGGGAAGT
AAGAATTTTTTAATTGGCATGAGACGGTTCCTCAACTGTTAAATTAACCAACTTTGACCTGTCTTTAGAA
```
```
AAAGGCTTATTTGTATGATTTTGGGCTAACTCCCCGGGGACCATATTAAATGACAAAAATGCTCCTTTGG
GTGACACACCCTACAAAGTATTTGCTGTTACGAACATAAACGCCCACATTCTTAATATCTAATATTTTTG
ACCAGTGATGTTTTATGCTGTCATCTGAACCCTAGAGAAGCAGTGTCAGAGGAAACCTTGGTGTCACATG
TGTCTTAGCAAAAGGGTTACCATGATCGAGGGTCATGTGACCAAAAGATGCTCCAGAGAAGCTTGAGAAT
TTGTTTCAAGTTGGGAGGAGGGTTGGAGATACAAAAATCACTCTGCTCTACAGGACTCTTCAGCTGTCTA
```
```
10  TGCAAGAAATTCCGTTTTCTCTTTCAGCACCTGGAAAGACACAGCAGCCCACCGAGGCGATAGGTGATTC
ACTAAGCACAAGAGGAATGTTTTCTAAGCAAGGCGTCCCTTGCCTCTCAAACAAATGCCCTCCAAGTTTG
TTAGGGTTTCTATTCCTGCAACTTGTGGTATCAAAACCACTTCCTGGAATTGTCAAAGCACTGCCAAAAT
AAATGTTTTTCCCCCTTCTAAGAAAAAAAAAATGACAGTGCTCATATTTGACACTTGTGTATTGGACTCT
CTTTTGAATGAATAAAAAGGAAAAGGGGTTTGGTGTAATTCCTGATGGGGTGCGTGTTGTTTTTCATGCC
```
```
15  ATGGTTTGTGAATTTTAATTGTGGTTTCCCATTTCGTTGTTGTAACTGGGCAGAAATTAAAAAAGAAAAA
TCAATAAAAATACAAAGAAATGG
```

>DEX0367_053.nt.1
```
TGCCACTTACGCGTCGCTCTTCCTCGTTTGCCCCTCGTGTTCATGGGAGCTCGTTTTCTTTTCCTCTAGG
```
```
CAGAGAAGAGGCGATGGCGGCGATGGCATCTCTCGGCGCCCTGGCGCTGCTCCTGCTGTCCAGCCTCTCC
CGCTGCTCAGCCGAGGCCTGCCTGGAGCCCCAGATCACCCCTTCCTACTACACCACTTCTGACGCTGTCA
TTTCCACTGAGACCGTCTTCATTGTGGAGATCTCCCTGACATGCAAGAACAGGGTCAGAACATGGCTCT
CTATGCTGACGTCGGTGGAAAACAATTCCCTGTCACTCGAGGCCAGGATGTGGGGCGTTATCAGGTGTCC
```
```
25  TGGAGCCTGGACCACAAGAGCGCCCACGCAGGCACCTATGAGGTTAGATTCTTCGACGAGGAGTCCTACA
GCCTCCTCAGGAAGGCTCAGAGGAATAACGAGGACATTTCCATCATCCCGCCTCTGTTTACAGTCAGCGT
GGACCATCGGGTGAGTGGCCTGGTCCCTCCTCCTTTTTGGGGTTGTTGGGCTGAGTGAAGGTTATCCTCT
CCACAGCCCCAGCTCTGCTGCTGGGCCGTGATTGGCCAGCATGTCTTGGTTCCCCTGGCGGAAGGTGACC
AGGGCTGGCTGGTCTGCTCACCTGTACTCCCCTGAGCTGGCTTGTGATCTCCTTTTTTTTCAGGGCACTTG
```
```
30  GAACGGGCCCTGGGTGTCCACTGAGGTGCTGGCTGCGGCGATCGGCCTTGTGATCTACTACTTGGCCTTC
AGTGCGAAGAGCCACATCCAGGCCTGAGGGCGGCACCCCAGCCCTGCCCTTGCTTCCTTĆAATAAACATC
ACAGGACCTGGGACTGC
```

>DEX0367_053.nt.2
```
TGCCACTTACGCGTCGCTCTTCCTCGTTTGCCCCTCGTGTTCATGGGAGCTCGTTTTCTTTTCCTCTAGG
```
```
CAGAGAAGAGGCGATGGCGGCGATGGCATCTCTCGGCGCCCTGGCGCTGCTCCTGGCGCAGAGACCCTTTCTCT
CGCTGCTCAGGAAACGCTGGGGCCAGTCTTCACCTTCTGTCAACTGGCCGGGGGCAGAGACCCTTTCTCT
CATCTTGAAAGCCCCAGTGCATTGCCTGCCTGCAGCCCCCACCCCCGATGTCTCTGCTGAGGATGCCTTC
AGTAACTCGTCCAGGCCGTTTGTGGCTCTGAATTGAGGCTGGCGTGGCCCCCGGGCCTCCGTGTCATAGG
```
```
40  TCCAAGTGTTGCCGAGGCCTGCCTGGAGCCCCAGATCACCCCTTCCTACTACACCACTTCTGACGCTGTC
ATTTCCACTGAGACCGTCTTCATTGTGGAGATCTCCCTGACATGCAAGAACAGGGTCCAGAACATGGCTC
TCTATGCTGACGTCGGTGGAAAACAATTCCCTGTCACTCGAGGCCAGGATGTGGGGCGTTATCAGGTGTC
CTGGAGCCTGGACCACAAGAGCGCCCACGCAGGCACCTATGAGGTTAGATTCTTCGACGAGGAGTCCTAC
AGCCTCCTCAGGAAGGCTCAGAGGAATAACGAGGACATTTCCATCATCCCGCCTCTGTTTACAGTCAGCG
```
```
45  TGGACCATCGGGGCACTTGGAACGGGCCCTGGGTGTCCACTGAGGTGCTGGCTGCGGCGATCGGCCTTGT
GATCTACTACTTGGCCTTCAGTGCGAAGAGCCACATCCAGGCCTGAGGGCGGCACCCCAGCCCTGCCCTT
GCTTCCTTCAATAAACATCACAGGACCTGGGACTGC
```

>DEX0367_054.nt.1
```
CAGATCTCAATTATCTAATTGCAATTGCAACGAGAACCAAAGCAGGGGAGCAGAGACAAACAATTTCTGA
```
```
GGTAACCAGATGGCTTTATTAACTCAAGTTCTCACCTAAAATTGCCCTCAAGAATCCTGTGGGAATGGGT
TGCAGTGGTGTGGCCCTGGATTCACAACCGACAGAGCTTCTGAATTCTGAGTGATCTGTACACAAACACA
CCTCTGCCTGGGTTACACGCCTCCACGTTCCTCTATCCAGTTCCCGCACCCTTCCCCCCAGGCCCCATTC
TTCAAGGCTTCAGAGCAGCGCTCCTCCGGTTAAAAGGAAGTCTCAGCACAGAATCTTCAAACCTCCTCGG
AGGCCACCAAAGATCCCTAACGCCGCCATGGAGACGAAGCACCTGGGCGGGGCGGAGCGGGGCGCGCGG
```
```
55  GCCCACACCTGTGGAGAGGGCCGCGCCCCAACTGCAGCGCCGGGGCTGGGGGAGGGGAGCCTACTCACTC
CCCCAACTCCCGGGCGGTGACTCATCAACGAGCACCAGCGGCCAGAGGTGAGCAGTCCCGGGAAGGGGCC
GAGAGGCGGGGCCGCCAGGTCGGGCAGGTGTGCGCTCCGCCCCGCCGCGCGCACAGAGCGCTAGTCCTTC
GGCGAGCGAGCACCTTCGACGCGGTCCGGGGACCCCCTCGTCGCTGTCCTCCCGACGCGGACCCGCGTGC
CCCAGGCCTCĆGCGCTGCCCGGCCGGCTCCTCGTGTCCCACTCCCGGCGCACGCCCTCCĆGCGAGTCCCGG
```
```
60  GCCCCTCCCGCGCCCCTCTTCTCGGCGCGCGCGCAGCATGGCGCCCCCGCAGGTCCTCGCGTTCGGGCTT
```

```
CTGCTTGCCGCGGCGACGGCGACTTTTGCCGCAGCTCAGGAAGAATGTGTCTGTGAAAACTACAAGCTGG
CCGTAAACTGCTTTGTGAATAATAATCGTCAATGCCAGTGTACTTCAGTTGGTGCACAAAATACTGTCAT
TTGCTCAAAGCGTGAGTAAAATATCCTAATTACCTGTAAGCTTTATTTTGACTTAATACTTCTTTAATTG
ATGTGCCTTGAGTTGGAAAGAGTTTTATTGGCTTAAATCTGAATCATGTTACAAAGTAAGTGTGGGAACA
CATAAATTTCAAATAATCTTTGACCCTGGAACTTTAGAGTTAATTTTTTTTTTCCCGTAÄTCATGAAATC
AGTTATTTTTCAGTTTGGCATTAAGGTTTCTTTTTCAGTGGCTGCCAAATGTTTGGTGATGAAGGCAGAA
ATGAATGGCTCAAAACTTGGGAGAAGAGCAAAACCTGAAGGGGCCCTCCAGAACAATGATGGGCTTTATG
ATCCTGACTGCGATGAGAGCGGGCTCTTTAAGGCCAAGCAGTGCAACGGCACCTCCATGTGCTGGTGTGT
GAACACTGCTGGGGTCAGAAGAACAGACAAGGACACTGAAATAACCTGCTCTGAGCGAGTGAGAACCTAC
TGGATCATCATTGAACTAAAACACAAAGCAAGAGAAAAACCTTATGATAGTAAAAGTTTGCGGACTGCAC
TTCAGAAGGAGATCACAACGCGTTATCAACTGGATCCAAAATTTATCACGAGTATTTTGTATGAGAATAA
TGTTATCACTATTGATCTGGTTCAAAATTCTTCTCAAAAAACTCAGAATGATGTGGACATAGCTGATGTG
GCTTATTATTTTGAAAAAGATGTTAAAGGTGAATCCTTGTTTCATTCTAAGAAAATGGACCTGACAGTAA
ATGGGGAACAACTGGATCTGGATCCTGGTCAAACTTTAATTTATTATGTTGATGAAAAAGCACCTGAATT
CTCAATGCAGGGTCTAAAAGCTGGTGTTATTGCTGTTATTGTGGTTGTGGTGATAGCAGTTGTTGCTGGA
ATTGTTGTGCTGGTTATTTCCAGAAAGAAGAGAATGGCAAAGTATGAGAAGGCTGAGATAAAGGAGATGG
GTGAGATGCATAGGGAACTCAATGCATAACTATATAATTTGAAGATTATAGAAGAAGGGAAATAGCAAAT
GGACACAAATTACAAATGTGTGTGCGTGGACGAAGACATCTTTGAAGGTCATGAGTTTGTTAGTTTAAC
ATCATATATTTGTAATAGTGAAACCTGTACTCAAAATATAAGCAGCTTGAAACTGGCTTTACCAATCTTG
AAATTTGACCACAAGTGTCTTATATATGCAGATCTAATGTAAAATCCAGAACTTGGACTCCATCGTTAAA
ATTATTTATGTGTAACATTCAAATGTGTGCATTAAATATGCTTCCACAGTAAAATCTGAAAAACTGATTT
GTGATTGAAAGCTGCCTTTCTATTTACTTGAGTCTTGTACATACATACTTTTTTATGAGCTATGAAATAA
AACATTTTAAACTGAA


>DEX0367_055.nt.1
TCGCCCTTTCGAGCGGCGCGCCCGGGCAGGTACTCCGTGTGGATCGGCGGCTCCATCCTGGCCTCGCTGTCC
ACCTTCCAGCÀGATGTGGATCAGCAAGCAGGAGTATGACGAGTCCGGCCCCTCCATCGTCCACCGCAAAT
GCTTCTAGGCGGACTATGACTTAGTTGCGTTACACCCTTTCTTGACAAAACCTAACTTGCGCAGAAAACA
AGATGAGATTGGCATGGCTTTATTTGTTTTTTTTTGTTTTGTTTTGGTTTTTTTTTTTTTTTTGGCTTGAC
TCAGGATTTAAAAAACTGGAACGGTGAAGGTGACAGCAGTCGGTTGGAGCGAGCATCCCCCAAAGTTCACA
ATGTGGCCGAGGACTTTGATTGCACATTGTTGTTTTTTTAATAGTCATTCCAAATATGAGATGCGTTGTT
ACAGGAAGTCCCTTGCCATCCTAAAAGCCACCCCACTTCTCTCTAAGGAGAATGGCCCAGTCCTCTCCCA
AGTCCACACAGGGGAGGTGATAGCATTGCTTTCGTGTAAATTATGTAATGCAAAATTTTTTTAATCTTCG
CCTTAATACTTTTTTATTTTGTTTTATTTTGAATGATGAGCCTTCGTGCCCCCCCTTCCCCCTTTTTTGT
CCCCCAACTTGAGATGTATGAAGGCTTTTGGTCTCCCTGGGAGTGGGTGGAGGCAGCCAGGGCTTACCTG
TACACTGACTTGAGACCAGTTGAATAAAAGTGCACACCTTAAAAATGA


>DEX0367_001.aa.1
MGRQKELVSRCGEMLHIRYRLLRQALAECLGTLILVMFGCGSVAQVVLSRGTHGGFLTINLAFGFAVTLG
ILIAGQVSGAHLNPAVTFAMCFLAREPWIKLPIYTLAQTLGAFLGAGIVFGLYYDAIWHFADNQLFVSGP
NGTAGIFATYPSGHLDMINGFFDQFIGTASLIVCVLAIVDPYNNPVPRGLEAFTVGLVVLVIGTSMGFNS
GYAVNPARDFGPRLFTALAGWGSAVFTTGQHWWWVPIVSPLLGSIAGVFVYQLMIGCHLEQPPPSNEEEN
VKLAHVKHKEQI


>DEX0367_001.aa.2
MGRQKELVSRCGEMLHIRYRLLRQALAECLGTLILVMFGCGSVAQVVLSRGTHGGFLTINLAFGFAVTLG
ILIAGQVSGAHLNPAVTFAMCFLAREPWIKLPIYTLAQTLGAFLGAGIVFGLYYDAIWHFADNQLFVSGP
NGTAGIFATYPSGHLDMINGFFDQVWAGDV


>DEX0367_002.aa.1
MEVKGQLISSPTFNAPAALFGEAAPQVKSERLRGLLDRQRTLQEALSLKLQELRKVCLQEAELTGQLPPE
CPLEPGERPQLVRRRPPTARAYPPPHPNQAHHSLCPAEELALEALEREVSVQQQIAAAARRLALAPDLST
EQRRRRRQVQADALRRLHELEEQLRDVRARLGLPVLPLPQPLPLSTGSVITTQGVCLGMRLAQLSQGEHP
LVRVGEWTLANGRGRAGMGDWPVKTGR


>DEX0367_003.aa.1
MVDVKCLSDCKLQNQLEKLGFSPGPILPSTRKLYEKKLVQLLVSPPCAPPVMNGPRELDGAQDSDDSEEL
NIILQGNIILSTEKSKKLKKWPEASTTKRKAVDTYCLDYKPSKGRRWAARAPSTRITYGTITKERDYCAE
DQTIESWREEGFPVGLKLAVLGIFIIVVFVYLTVENKSLFG
```

>DEX0367_004.aa.1
MPFQKGMPFDLCFLVQSSDFKVMVNGILFVQYFHRVPFHRVDTISVNGSVQLSYISFQPPGVWPANPAPI
TQTVIHTVQSAPGQMFSTPAIPPMMYPHPAYPMPFITTILGGLYPSKSILLSGTVLPSAQRFHINLCSGN
HIAFHLNPRFDENAVVRNTQIDNSWGSEERSLPRKMPFVRGQSFSVWILCEAHCLKVAVDGQHLFEYYHR
LRNLPTINRLEVGGDIQLTHVQT

>DEX0367_005.aa.1
MEKSGRRWLASAAPPLGRLRRRESGAEQGGLSVRATRVSLVRSALDCAPRSGVRRPGSCFCRCRRRIPVA
RRARLPQACSQHRTEPSGGRGWSARPAWERQGRRCNLLTAKKPGE

>DEX0367_006.aa.1
MRHSISNLKKKKKKKKKTSGKKVRGILSLKLVSEGTGEEKTTVPNEKRTGNLILIGMHQILLCTFASSIS
RRIVQNVYFLPMLRKQVYRTYSGLIASEWQIRIGIQSPCCGLLQQENQATQMILFSLFGFVKCHLVLFPS
NIEEVVGLKLWDLHYAYTFLFMPLFREADYCFFKMMHWRRCESKIATWHYLPRINEKGKKTIFSFFKHFS
EKIQLPFLIGERHHARLIFAFLVETGFHHVGQDGLDLLISCSAHLGLLSAGITGMSHCARSTILFSVSHP
YQIIEPSVCMFLNLVYNSTSLTYKSTISYEFFIEVGSILKWTENLIPGRARWLMPVIPYFGRPRWVDHLR
LGVRDQPGQHGETPSLLKNTKISQAWWLSVIPANGEAEAQESLEPEEAEVAVSRDHTTALHPGQWSETLS
QKIINEDMIPACFIQLERHTTHEIIGEHVNVYLLVQLRKREEYVLVSKVLNKTEVASTVAHVFFGLTFFF
SSTFCNFYDLGHEVLPLRHNQYPSRKGLLIPGVKIPSLRGSHYGSPGVKIPSSQESYLQHLGKIPEGATS
NRKMKERFNFSTQVTNPMHSIVYVICRKGTGGVRALWVTALLVVTFTLLFLLNLWFSRLKNEIIGQINSS
QPFLEQQIRLSLKSFSKISCFQSLPVIAIIPRLIKSVLVNQFLYFFFLSFFYSVFRYHLTNNLLLILLR

>DEX0367_006.aa.2
MGRLRHKNLLNLGGGGCSEPRSHHCTPSWAMERDSVSKNNKEDMIPACFIQLERHTTHEIIGEHVNVYLL
VQLRKREEYVLVSKVLNKTEVASTVAHVFFGLTFFFSSTFCNFYDLGHEVLPLRHNQYPSRKGLLIPGVK
IPSLRGSHYVIPRSKDSYLAGILLAHLGKIPEGATSNRKMKERFNFSTQVTNPMHSIVYVICRKGTGGVR
ALWVTALLVVTFTLLFLLNLWFSRLKNEIIGQINSSQPFLEQQIRLSLKSFSKISCFQSLPVIAIIPRLI
KSVLVNQFLYFFFLSFFYSVFQVSFDQQLALNLTAMTEISKEESIVVIGLVIIILKTIHYFLKLNAVCLI
SVEYDKNVNEKPNMTLTLALNFIFTCVLKVLYWPVGKLICIRFTPGFGRNSFLKIQLADLKMFFIPQNVS
LLCHSYRLSAFFDHHVNYAVVNHGVQADELVVFVKQNVMSIGRDSSSGEHGSFPVIPALMNRLKTAVYYG
QFNFTGLPGLAFQLLSCRETWCCTVHILEKWQECSLELQITQQRIPYQYIGFSCYIEIWYFSDGYGFCLT
CVSVLLERQYRIKNFLMSFLKIEIIMALLVVLCSDRPLIKKSFNPAFHFTSPPFIFNIHSLKIVMIFSVI
GYVKKFHFKVLICNNLHFLLNWRTFLRQTYFYELIFSLVKENSLVSYREWLALCIPHPLCGLPVAVVLLK
IADILINMMIFGTGLSLLLISNKVGQSNLNYFNKHNLAFLYMRKYFQKITRFISRIRDGKYQ

>DEX0367_007.aa.1
MLANDVRHQQEMWGFRKVEGGVVQSLGKSSVEGETDGTISEFREIQRLAAFASFLSHAPPLNARRLLTPP
PRRRPRCTPAAAMADVSERTLQLSVLVAFASGVLLGWQANRLRRRYLDWRKRRLQDKLAATQKKLDLA

>DEX0367_008.aa.1
MPCARAGGLGLQTPNLNGHPRAEPPEGTGGFHFQTGILAASLLPPAEEETLLYILTFCRQVKRRTQTFGD
ERETKTEPSVQVTASQSRHLTDPTYCLFLNMNDCRSLPETVSEKTATSYFLYVFPIKRLSWVGLRITEGK
RSQFQVTQAIFCVKREQDKSQPQQQNPKPPLRLLWQSNT

>DEX0367_009.aa.1
ISVLTWAVFTPPLPSRYFSCAHSTDREAEAGEVRTRLRSYGLPWDLAEDGGRAGPSGLETLTPYSPTPSF
TWSDARLHRGLVTLLTGEIVDAFSLEFRTLYAASCPLPPAPPQKPSVIGGLQRGRSPHRVSRRRSVAPAS
PPPPDGPLAHRLAACRVSPATPGPALSDILRSVQRARTPSGPPARPSRSMWDLSRLSQLSGSSDGDNELK
KSWGSKDTPAKALMRQRGTGGGPWGEVDSRPPWGGALPLPPAHRLRYLSPARRRFGGDATFKLQEPRGVR
PSDWAPRAGLGGQP

>DEX0367_010.aa.1
MALQLSREQGITLRGSAEIVAEFFSFGINSILYQRGIYPSETFTRVQKYGLTLLVTTDLELIKYLNNVVE
QLKVHPEKSLRKLSRMKSVQ

>DEX0367_011.aa.1
RTVTVRTRIAVLSLRPQCGGILFRHVVVLTLGNGLGQNLDLASVQAHAAVQGRRVLIPGVNIRQENLGRG
RFHDHVQDAAVGGVGQALRCHQHKAVGLTQHLEPFPDLRAECRVAEHQPGFVQDDERPPVWWNSNPEKDI
FVVRENGTTCLMAEFAAKFIVPYDVWASNYVDLITEQADIALTRGAEVKGRCGHSQSELQVFWVDRAYAL
KMLFVKESHNMSKGPEATWRLSKVQFVYDSSEKTHFKDAVSAGKHTANSHHLSALVTPAGKSYECQAQQT

ISLASSDPQKTVTMILSAVHIQPFDIISDFVFSEEHKCPVDEREQLEETLPLILGLILGLVIMVTLAIYH
VHHKMTANQVQIPRDRSQYKHMG

>DEX0367_012.aa.1
MVLSSPLALAGDTQPRFLWQDKYGVSFLQRDGAGAVSWKECIYNQEEFVRFDSDVGEYRAVTELGRPVAD
PS

>DEX0367_012.aa.2
MVLSSPLALAGDTQLHPEVTVYPAKTQPLQHHNLLVCSVSGFYPGSIEVRWFRNGQEEKAGVVSTGLIQN
GDWTFQTLVMLETVPRSGEVYTCQGGAPKCDEPSHSGMESTV

>DEX0367_013.aa.1
MMPPYPGPGHCHQQLNSLPTMEGPPTFNPPVPYFGRLQGGLTARRTIIIKGYVPPTGKSFAINFKVGSSG
DIALHINPRMGNGTVVRNSLLNGSWGSEEKKITHNPFGPGQFFDLSIRCGLDRFKVYANGQHLFDFAHRL
SAFQRVDTLEIQGDVTLSYVQI

>DEX0367_013.aa.2
GFRGPAHWPLPTGLHLTPPSLSVPSFAINFKVGSSGDIALHINPRMGNGTVVRNSLLNGSWGSEEKKITH
NPFGPGQFFDLSIRCGLDRFKVYANGQHLFDFAHRLSAFQRVDTLEIQGDVTLSYVQI

>DEX0367_014.aa.1
MAATCEISNIFSNYFSAMYSSEDSTLASVPPAATFGADDLVLTLSNPQMSLEGTEKASWLGEQPQFWSKT
QVLDWISYQVEKNKYDASAIDFSRCDMDGATLCNCALEELRLVFGPLGDQLHAQLRDLTSSSSDELSWII
ELLEKDGMAFQEALDPGPFDQGSPFAQELLDDGQQASPYHPGSCGAGAPSPGSSDVSTAGTGASRSSHSS
DSGGSDVDLDPTDGKLFPSGESREVPKRASHLAMHRGPGSSCSLFL

>DEX0367_014.aa.2
MAATCEISNIFSNYFSAMYSSEDSTLASVPPAATFGADDLVLTLSNPQMSLEGTEKASWLGEQPQFWSKT
QVLDWISYQVEKNKYDASAIDFSRCDMDGATLCNCALEELRLVFGPLGDQLHAQLRDLTSSSSDELSWII
ELLEKDGMAFQEALDPGPFDQGSPFAQELLDDGQQASPYHPGSCGAGAPSPGSSDVSTAGTGTGWEVCPE
SRSVVEQRVGRQGTYSDPAPRTGASRSSHSSDSGGSDVDLDPTDGKLFPSDGFRDCKKGDPKHGKRKRGR
PRKLSKEYWDCLEGKKSKHAPRGTHLWEFIRDILIHPELNEGLMKWENRHEGVFKFLRSEAVAQLWGQKK
KNSNMTYEKLSRAMRYYYKREILERVDGRRLVYKFGKNSSGWKEEEVLQSRN

>DEX0367_015.aa.1
MSLPVKPELLGDLEIPAVPILHSMVQKFPGVSFGISTDSEVLTHYNITGNTICLFRLVDNEQLNLEDEDI
ESIDATKLSRFIEINSLHMVTEYNPVTVIGLFNSVIQIHLLLIMNKASPEYEENMHRYQKAAKLFQGKIL
FILVDSGMKENGKVISFFKLKESQLPALAIYQTLDDEWDTLPTAEVSVEHVQNFCDGFLSGKLLKENRES
EGKTPKVEL

>DEX0367_015.aa.2
MKETCQLEIQVDNEQLNLEDEDIESIDATKLSRFIEINSLHMVTEYNPVTVIGLFNSVIQIHLLLIMNKA
SPEYEENMHRYQKAAKLFQGKILFILVDSGMKENGKVISFFKLKESQLPALAIYQTLDDEWDTLPTAEVS
VEHVQNFCDGFLSGKLLKENRESEGKTPKVEL

>DEX0367_016.aa.1
MVCLFLKCNLQNSPERNNSFWEVVAAAGTVAPWEIVKNPEFILGGATRTDICQGELGDCWLLAAIASLTL
NQKALARVIPQDQSFGPGYAGIFHFQFWQHSEWLDVVIDDRLPTFRDRLVFLHSADHNEFWSALLEKAYA
KLNGSYEALKGGSAIEAMEDFTGGVAETFQTKEAPENFYEILEKALKRGSLLGCFIDTRSAAESEARTPF
GLIKGHAYSVTGIDQVSFRGQRIELIRIRNPWGQVEWNGSWSDRMAFKDFKAHFDKVEICNLTPDALEED
ADPQIKLSLTEKDEGQEECSFLVALMQKDRRKLKRFGANVLTIGYAIYECPDKDEHLNKDFFRYHASRAR
SKTFINLREVSDRFKLPPGEYILIPSTFEPHQEADFCLRIFSEKKAITRDMDGNVDIDLPEPPKPTPPDQ
ETEEEQRFRALFEQVAGEDMEVTAEELEYVLNAVLQKKKDIKFKKLSLISCKNIISLMDTSGNGKLEFDE
FKVFWDKLKQWINLFLRFDADKSGTMSTYELRTALKAAGFQLSSHLLQLIVLRYADEELQLDFDDFLNCL
VRLENASRVFQALSTKNKEFIHLNINEFIHLTMNI

>DEX0367_016.aa.2
MAFKDFKAHFDKVEICNLTPDALEEDAIHKWEVTVHQGSWVRGSTAGGCRNFLDTFWTNPQIKLSLTEKD
EGQEECSFLVALMQKDRRKLKRFGANVLTIGYAIYECPDKDEHLNKDFFRYHASRARSKTFINLREVSDR
FKLPPGEYILIPSTFEPHQEADFCLRIFSEKKAITRDMDGNVDIDLPEPPKPTPPDQETEEEQRFRALFE

QVAGEDMEVTAEELEYVLNAVLQKKKDIKFKKLSLISCKNIISLMDTSGNGKLEFDEFKVFWDKLKQWIN
LFLRFDADKSGTMSTYELRTALKAAGFQLSSHLLQLIVLRYADEELQLDFDDFLNCLVRLENASRVFQAL
STKNKEFIHLNINEFIHLTMNI

5 >DEX0367_017.aa.1
ALSVVAAEVRVVGRNGESSELDLQGIRIDSDISGTLKFACESIVEEYEDELIEFFSREADNVKDKLCSKR
TDLCDHALHISHDEL

>DEX0367_018.aa.1
10 MADYSTVPPPSSGSAGGGGGGGGGGGVNDAFKDALQRARQIAAKIGGDAGTSLNSNDYGYGGQKRPLEDG
DQPDAKKVAPQNDSFGTQLPPMHQQQSRSVMTEEYKVPDGMVGFIIGRGGEQISRIQQESGCKIQIAPDS
GGLPERSCMLTGTPESVQSAKRLLDQIVEKGRPAPGFHHGDGPGNAVQEIMIPASKAGLVIGKGGETIKQ
LQERAGVKMVMIQDGVAAAAAVFFTSIIRLAFNKEKVLTPVLNCT

15 >DEX0367_019.aa.1
MAEHFLTLLVVPAIKKDYGSQEDFTQVWNTTMKGLKCCGFTNYTDFEDSPYFKENSAFPPFCCNDNVTNT
ANETCTKQKAHDQKVEGCFNQLLYDIRTNAVTVGGVAAGIGGLELAAMIVSMYLYLPGRPAWSRPRYSDS
VGRVSFPSLVCFLMRLEAMAKTFRNSLRMEKDST

20 >DEX0367_019.aa.2
MQCFSFIKTMMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGPLSSSAMQFVNVGYFLIAAGVVVFAL
GFLGCYGAKTESKCALVTFFFILLLIFIAEVAAAVVALVYTTMAEHFLTLLVVPAIKKDYGSQEDFTQVW
NTTMKGLKCCGFTNYTDFEDSPYFKENSAFPPFCCNDNVTNTANETCTKQKAHDQKVEGCFNQLLYDIRT
NAVTVGGVAAGIGGLEFFSNSARRPPLPESLYSTPIRRDHVFLQPSPP

25 >DEX0367_020.aa.1
MLGAALRRCAVAATTRADPRGLLHSARTPGPAVAIQSVRCYSHGSQETDEEFDARWVTYFNKPDIDAWEL
RKGINTLVTYDMVPEPTVRILEVVKDKAGPHKEIYPYVIQELRPTLNELGISTPEELGLDKV

30 >DEX0367_021.aa.1
MGQDWGPHHMSREDLICGGGKGSLVWPTWSTTSLWPQVLLVQFKGTGKYYAIKALKKQEVLSRDEIESLY
CEKRILEAVGCTGHPFLLSLLACFQTSSHACFVTEFVPGGDLMMQIHEDVFPEPQARFYVACVVLGLQFL
HEKKIIYRDLKLDNLLLDAQGFLKIADFGLCKEGIGFGDRTSTFCGTPEFLAPEVLTQEAYTRAVDWWGL
GVLLYEMLVGECPFPGDTEEEVFDCIVNMDAPYPGFLSVQGLEFIQKLLQKCPEKRLGAGEQDAEEIKVQ
35 PFFRTTNWQALLARTIQPPFVPTLCGPADLRYFEGEFTGLPPALTPPAPHSLLTARQQAAFRDFDFVSER
FLEP

>DEX0367_021.aa.2
MEEGAPRQPGPSQWPPEDEKEVIRRAIQKELKIKEGVENLRRVATDRRHLGHVQQLLRSSNRRLEQLHGE
40 LRELHARILLPGPGPGPAEPVASGPRPWAEQLRARHLEALRRQLHVELKVKQGAENMTHTCASGTPKERK
LLAAAQQMLRDSQLKVALLRMKISSLEASGSPEPGPELLAEELQHRLHVEAAVAEGAKNVVKLLSSRRTQ
DRKALAEAQAQLQESSQKLDLLRLALEQLLEQLPPAHPLRSRVTRELRAAVPGYPQPSGTPVKPTALTGT
LQVRLLGCEQLLTAVPGRSPAAALASSPSEGWLRTKAKHQRGRGELASEVLAVLKVDNRVVGQTGWGQVA
EQSWDQTFVIPLERARELEIGVHWRDWRQLCGVAFLRLEDFLDNACHQLSLSLVPQGLLFAQVTFCDPVI
45 ERRPRLQRQERIFSKRRGRLPGLRLCVRAIPGTLRASPGTSVPFPHRLLEPLLVHPCALPGGPGLAGYF

>DEX0367_022.aa.1
MDMAWQMMQLLLLALVTAAGSAQPRSARARTDLLNVCMNAKHHKTQPSPEDELYGQVGAPQGPSPGSVPL
DDLPGAEEPEYGGDGCGGERLSPVSSPPSAVPGRRMPAARPAPARSCTRTPPACTTLTGITVVRWNPPAS
50 ATLSRTAVSECSPNLGPWIRQVNQSWRKERILNVPLCKEDCERWWEDCRTSYTCKSNWHKGWNWTSGINE
CPAGALCSTFESYFPTPAALCEGLWSHSFKVSNYSRGSGRCIQMWFDSAQGNPNEEVAKFYAAAMNAGAP
SRGIIDS

>DEX0367_023.aa.1
55 MTHSLVCPETVSRVSSVLNRNTRQFGKKHLFDQDEETCWNSDQGPSQWVTLEFPQLIRVSQLQIQFQGGF
SSRRGCLEGSQGTQALHKIVDFYPEDNNSLQDILPLLLGWVVCASGSFGLTQTRMGTLMCDLEHVTSRTP
LFL

>DEX0367_023.aa.2
60 MLELRPGSQGTQALHKIVDFYPEDNNSLQTFPIPAAEVDRLKVTFEDATDFFGRVVIYHLRVLGEKV

>DEX0367_024.aa.1
QVPHASGCIGEDPQGIPLPLDEAHPTGGCTNPYGKSKFFIEEMIRDLCQADKGWTAALGLDRMCEDLWRW
QKQNPSGFGTQA

>DEX0367_024.aa.2
MAEKVLVTGGAGYIGSHTVLELLEAGYLPVVIDNFHNAFRGGGSLPESLRRVQELTGRSVEFEEMDILDQ
GALQRLFKKYSFMAVIHFAGLKAVGESVQKPLDYYRVNLTGTIQLLEIMKAHGVKNLVFSSSATVYGNPQ
YLPLDEAHPTGGCTNPYGKSKFFIEEMIRDLCQADKTWNAVLLRYFNPTGAHASGCIGEDPQGIPNNLMP
YVSQVAIGRREALNVFGNDYDTEDGTGVRDYIHVVDLAKGHIAALRKLKEQCGCRVGREGRSEGGEGPDP
GRAAQRRGQSSPLHKPCSPWARSTTWARAQAIQCCRWSRLWRRPLGRRSRTRWWHGGKVMWQPVTPTPAW
PKRSWGGQQP

>DEX0367_024.aa.3
MAEKVLVTGGAGYIGSHTVLELLEAGYLPVVIDNFHNAFRGGGSLPESLRRVQELTGRSVEFEEMDILDQ
GALQRLFKKYSFMAVIHFAGLKAVGESVQKPLDYYRVNLTGTIQLLEIMKAHGVKNLVFSSSATVYGNPQ
YLPLDEAHPTGGCTNPYGKSKFFIEEMIRDLCQADKTWNAVLLRYFNPTGAHASGCIGEDPQGIPNNLMP
YVSQVAIGRREALNVFGNDYDTEDGTGVRDYIHVVDLAKGHIAALRKLKEQCGCRVGREGRSEGGEGPDP
GRAAQRRGQSSPLHKPCSPWARSTTWARAQAIQCCRWSRLWRRPLGRRSGPPTPPTSPTSPHPALSNRAA
LALPTHLSGGYLALPSLLSLPSTRCLRASRCSALLLLKDLASSWARAGSAKLQLPPVLQIPYKVVARREG
DVAACYANPSLAQEELGWTAALGLDRMCEDLWRWQKQNPSGFGTQA

>DEX0367_024.aa.4
MAEKVLVTGGAGYIGSHTVLELLEAGYLPVVIDNFHNAFRGGGSLPESLRRVQELTGRSVEFEEMDILDQ
GALQRLFKKYSFMAVIHFAGLKAVGESVQKPLDYYRVNLTGTIQLLEIMKAHGVKNLVFSSSATVYGNPQ
YLPLDEAHPTGGCTNPYGKSKFFIEEMIRDLCQADKTWNAVLLRYFNPTGAHASGCIGEDPQGIPNNLMP
YVSQVAIGRREALNVFGNDYDTEDGTGVRDYIHVVDLAKGHIAALRKLKEQCGCRVGREGRSEGGEGPDP
GRAAQRRGQSSPLHKPCSPWARSTTWARAQAIQCCRWSRLWRRPLGRR

>DEX0367_025.aa.1
MASTSYDYCRVPMEDGDKRCKLLLGIGILVLLIIVILGVPLIIFTIKANSEACRDGLRAVMECRNVTHLL
QQELTEAQKGFQDVEAQAATCNHTVMALMASLDAEKAQGQKKVEELEGEITTLNHKLQDASAEVERLRRE
NQVLSVRIADKKYYPSSQDSSSAAAPQQHNQQLGRRSC

>DEX0367_025.aa.2
MASTSYDYCRVPMEDGDKRCKLLLGIGILVLLIIVILGVPLIIFTIKANSEACRDGLRAVMECRNVTHLL
QQELTEAQKGFQDVEAQAATCNHTVMALMASLDAEKAQGQKKVEELEVTTLNHKLQDASRPGFLFSVAPP
LQTRLRRENQVLSVRIADKKYYPSSQDSSSAAAPQLLIVLLGLSALLQ

>DEX0367_025.aa.3
MPWIGGGLIAVWGGAINGGGSIDGGEGLMDGWGGSSDGGGAWGRAWSGASACCLQSALSPSSGPLGTSWK
VRPARLFA

>DEX0367_026.aa.1
MSPLSPGSPLPPLARADLTAILTGCPPLSACLVLAPRPHRRARLLPSEGLLTLRAKPPSEAEYTDVLQKI
KYAFSLLARLRGNIADPSSPELLHFLFGPLQMIVNTSGGPEFASSVRRPHLTSDAVALLRDNVTPRENEL
WTSLGDSWTRPGLELSPEEGPPYRPEFFSGWEPPVTDPQSRAWEDPVEKQLQHERRRRQQSAPQVAVNGH
RDLEPESEPQLESETAGKWVLCNYDFQARNSSELSVKQRDVLEVLDDSRKWWKVRDPAGQEGYVPYNILT
PYPGPRLHHSQSPARSLNSTPPPPPAPAPAPPPALARPRWDRPRWDSCDSLNGLDPSEKEKFSQMLIVNE
ELQARLAQGRSGPSRAVPGPRAPEPQLSPGSDASEVRAWLQAKGFSSGTVDALGVLTGAQLFSLQKEELR
AVSPEEGARVYSQVTVQRSLLEDKEKVSELEAVMEKQKKKVEGEVEMEVI

>DEX0367_027.aa.1
MLWVWFHITAIKLQREEEEAFASSQSSQGAQSLIFSKFEGKKTNKKTRKVTTVKKSSVRLPGSDQRRILK
WIPGVCLETSWPASPSAASTSSMPSRSPRTSWRQQMTSSPPSSTLF

>DEX0367_028.aa.1
MATKGGTVKAASGFNAMEDAQTLRKAMKGLAQTTRKQKIHQTFMNSQLLERQGTNLKNQNKVSLLKLLQE
TGTDEDAIISVLAYRNTAQRQEIRTAYKSTIGRDLIDDLKSELSGNFEQVIVGMMTPTVLYDVQELRRAM
KGAGTDEGCLIEILASRTPEEIRRISQTYQQQYGRSLEDDIRSDTSFMFQRVLVSLSAGGRDEGNYLDDA

LVRQDAQDLYEAGEKKWGTDEVKFLTVLCSRNRNHLLHVFDEYKRISQKDIEQSIKSETSGSFEDALLAI
VKCMRNKSAYFAEKLYKSMKGLGTDDNTLIRVMVSRAEIDMLDIRAHFKRLYGKSLYSFIKGDTSGDYRK
VLLVLCGGDD

>DEX0367_028.aa.2
MYTLELFFICFSHPQGGRDEGNYLDDALVRQDAQDLYEAGEKKWGTDEVKFLTVLCSRNRNHLLHVFDEY
KRISQKDIEQSIKSETSGSFEDALLAIVKCMRNKSAYFAEKLYKSMKGLGTDDNTLIRVMVSRAEIDMLD
IRAHFKRLYGKSLYSFIKGDTSGDYRKVLLVLCGGDD

>DEX0367_029.aa.1
MYFDWGPGEMLVCETSFNKKEKSEMVPSCPFIYIIRKDVDVYSQILRKLFNESHGIFLGLQRIDEELTGK
SRKSQLVRVSKNYRSVIRACMEEMHQVAIAAKDPANGRQFSSQVSILSAMELIWNLCEILFIEVAPAGPL
LLHLLDWVRLHVCEVDSLSADVLGSENPSKHDSFWNLVTILVLQGRLDEARQMLSKEADASPASAGICRI
MGDLMRTMPILSPGNTQTLTELELKWQHWHEECERYLQDSTFATSPHLESLLKIMLGDEAALLEQKELLS
NWYHFLVTRLLYSNPTVKPIDLHYYAQSSLDLFLGGESSPEPLDNILLAAFEFDIHQVIKECRNKTDLSR
RSLLDAGSIKGESILLFPVAEEKEKYHEEG

>DEX0367_029.aa.2
MYFDWGPGEMLVCETSFNKKEKSEMVPSCPFIYIIRKDVDVYSQILRKLFNESHGIFLGLQRIDEELTGK
SRKSQLVRVSKNYRSVIRACMEEMHQVAIAAKDPANGRQFSSQVSILSAMELIWNLCEILFIEVAPAGPL
LLHLLDWVRLHVCEVDSLSADVLGSENPSKHDSFWNLVTILVLQGRLDEARQMLSKEADASPASAGICRI
MGDLMRTMPILSPGNTQTLTELELKWQHWHEECERYLQDSTFATSPHLESLLKIMLGDEAALLEQKELLS
NWYHFLVTRLLYSNPTVKPIDLHYYAQSSLDLFLGGESSPEPLDNILLAAFEFDIHQVIKECSFLLKTGQ
FLAVWQEETAGVHFTGSWARCRQFPGALQVLQKYRAKSIALSNWWFVAHLTDLLDHCKLLQSHNLYFGSN
MREFLLLEYASGLFAHPSLWQLGVDYFDYCPELGRVSLELHIERIPLNTEQKALKVLRICEQRQMTEQVR
SICKILAMKAVRNNRLGSALSWSIRAKDAAFATLVSDRFLRDYCERGCFSDLDLIDNLGPAMMLSDRLTF
LGKYREFHRMYGEKRFADAASLLLSLMTSRIAPRSFWMTLLTDALPLLEQKQVIFSAEQTYELMRCLEDL
TSRRPVHGESDTEQLQDDDIETTKVEMLRLSLARNLARAIIREGSLEGS

>DEX0367_030.aa.1
RCNTPTSSQIKFQHILYASVTKQPHNLVFIVDLYRMKEENPLWHANEIFLVGTADEATRAEIQIRIEANE
ALVKALE

>DEX0367_031.aa.1
MQERNGCCEIPLRNYIIQVMKITQFYFQRRTDEKGL

>DEX0367_032.aa.1
MLPAVGSADEEEDPAEEDCPELVPIETTQSEEEEKSGLGAKIPVTIITGYLGAGKTTLLNYILTEQHSKR
VAVILNEFGEGSALEKSLAVSQGGELYEEWLELRNGCLCCSVKDSGLRAIENLMQKKGKFDYILLETTGL
ADPGAVASMFWVDAELGSDIYLDGIITIVDSKYGLKHLTEEKPDGLINEATRQVALAEIGPPPNFFXPFS
XKFFF

>DEX0367_033.aa.1
MSHRQVHDDLNKLLKIMLINSFGSVIIFVFINILSQFSSFIFISEISMSWNKSCVLISLLCNNLVCLTFL
TFISNICFIKNNKHAVIDFSYFKWMSEQVTKIFCEFFSVWCLPMHLRIWGLSEIVLPCYGTEVGLESFSM
KIRCPEYEFLLLGPESYIKYSLKFLEATAPSLSSVIFWAYVKIITQSPVFINCFFIFKPNLMLIVICYLF
SPDLNHWIQLNEFELSLNNSKRNNVYSDGGNFLSTCSPILNEVKSNHVTIRVLEKLNILYIGYLTPHFYI
TCYIKGGGIKEIQKLQRYLECTYLLILFVISLFHLLSNKLLEKFLFFSCFFSYKNVFEKLINFRIEKIIE
SLKKTYFIEVITKIIFNLDSTVIQILHHLPTSMNFMYKFFKSQSFFFLINWMYFTEFPTAHVSFLPFRVD
LSNVLDLHAFDSLSGISLQKKLQHVPGTQPHLDQSIVTITFDVPGNAKEEHLNMFIQNLLWEKNVRNKDN
HCMEVIRLKVQFTVADFWTKSFSWLLEKLYLVLNRNTGFSTNHLCLLSFFFIIFMTEKELWKSLHKAGFI
CTTFFRVAARTNLCALKCYLLLSVPKYREIMLQISLLLNIMLPDAFEQTLNICCTLNKVQRTRRILVLYL
ETHSHYLIFGYLSHERYFFYGSSDSQSVCLTSQLSVYSCVFTSVHKVFGEIKNIISNEINFIPIGASLSD
NSFLISANQYTMSSYSDKYNSFSLFQHCSLIATHFYNKLFNITNSFNFSTFPTKTVKHYIKSLSIGYDTY
FIILFQVLVVINNTEKPSIIYVLTLSLEKGIVQKKINTQKPFLKIKNIKKLLVIHKYLELSNFLSFKSLY
FLSEYQYINPLTLMLISAFKPELRLINVQSILLGAGLVSIKDKSQQVIVQGVHELYDLEETPVSWKDDTE
RTNRLVLIGRNLDKDILKQLFIATVTETEKQWTTHFKEDQVCT

>DEX0367_033.aa.2
MRLSNRQPGALRLTAGSLVPLSLYLRNSFFGSTAEALGEWLCLLWQRLEVLTDCHKYYAVTAAAAYMHVN

SWGINLVCILRSHSSAGRGSRRMPFSVSPLQPYTKCAPCVSNSIVEVSDNLTYTMSHSSVSVLFLLVFYN
SFLLNFSPLYKMSHRQVHDTYNKLLKIMLINSFGSVIIFVFINILSQFSSFIFISEISMSWNKSCVLISL
LCNNLVCLTPLTFISNICFIIEQKHAVIDFSYFKWMSEQVTKIFCEFFSVWCLPMHLRIQGLSEIVLPCY
GTEVGLESFSMKIRCPEYEFLLLGPESYIKYSLKFLEATAPSLSSVIFWAYVKIITQSPVFINCFFIFKP

5   NLMLIVICYLFSPDLNHWIQLNEFELSLNNSKRNNVYSDGGNFLSTCSPILNEVKSNHVTIRVLEKLNIL
YIGYLTPHFYITCYIKGGGIKEIQKLQRYLECTYLLILFVISLFHLLSNKLLEKFLFFSCFFSYKNVFEK
LINFRIEKIIESLKKTYFIEVITKIIFNLDSTVIQILHHLPTSMNFMYKIFKSQSFFFLINWMYFTEFPT
AHVSFLPFRVDLSNVLDLHAFDSLSGISLQKKLQHVPGTQPHLDQSIVTITFDVPGNAKEEHLNMFIQNL
LWEKNVRNKDNHCMEVIRLKGLVSIKDKSQQVIVQGVHELYDLEETPVSWKDDTERTNRLVLIGRNLDKD

10   ILKQLFIATVTETEKQWTTHFKEDQVCT .

>DEX0367_034.aa.1
MAEISDLDRQIEQLRRCELIKESEVKALCAKAREILVEESNVQRVDSPVTVCGDIHGQFYDLKELFRVGG
DVPETNYLFMGDFVDRGFYSVETFLLLLALKVRYPDRITLIRGNHESRQITQVYGFYDECLRKYGSVTVW

15   RYCTEIFDYLSLSAIIDGKIFCVHGGLSPSIQTLDQIRTIDRKQEVPHDGPMCDLLWSDPEDTTGWGVSP
RGAGYLFGSDVVAQFNAANDIDMICRAHQLVMEGYKWHFNETVLTVWSAPNYCYRCGNVAAILELDEHLQ
KDFIIFEAAPQETRGNPARPLPPPTLLALAPL

>DEX0367_034.aa.2
20   MQPLLHSHSGPPCDNTGCPSVSGSAPFTLMCCLDGDCMITVIMRGPQALGVLSPLELICPYICCPGPRHP
TDHIQAPCSGPTGLWWTWGGARLLTLKGQRWQPEKPEDIPLHPSLSASDTTGWGVSPRGAGYLFGSDVVA
QFNAANDIDMICRAHQLVMEGYKWHFNETVLTVWSAPNYCYR

>DEX0367_035.aa.1
25   MLLPWATSAPGLAWGPLVLGLFGLLAASQPQAVPPYASENQTCRDQEKEYYEPQHRICCSRCPPGTYVSA
KCSRIRDTVCATCAENSYNEHWNYLTICQLCRPCDPVMGLEEIAPCTSKRKTQCRCQPGMFCAAWALECT
HCELLSDCPPGTEAELKDEVGKGNNHCVPCKAGHFQNTSSPSARCQPHTRCENQGLVEAAPGTAQSDTTC
KNPLEPLPPPEMSEPALSKGVENLQALLYQAATGSSEASFPTLSPLYTPPQVQVQQGNPELLYSSPSVQWL
RPQKCGSSLCLFTTPSPTLPYCLPIPLPDLENQLPKLPSCTHKPAQSWSLSRRAPTPPPNMPIHDTVSPG
30   CQEVLKSNLVPLLYNPREVSLILPLGAALCLEGKKLLPFLCLGCPGIWKALPSPPPSALLGAVITLLSAV
LAGTMLMLAVLLPLAFFLLLATVFSCIWKSHPSLCRKLGSLLKRRPQGEGPNPVAGSWEPPKAHPYFPDL
VQPLLPISGDVSPVSTGLPAAPVLEAGVPQQQSPLDLTREPQLEPGEQSQVAHGTNGIHVTGGSMTITGN
IYIYNGPVLGGPPGPGDLPATPEPPYPIPEEGDPGPPGLSTPHQEDGKAWHLAETEHCGATPSNRGPRNQ
FITHD

35
>DEX0367_035.aa.2
MLLPWATSAPGLAWGPLVLGLFGLLAASQPQAVPPYASENQTCRDQEKEYYEPQHRICCSRCPPGTYVSA
KCSRIRDTVCATCAENSYNEHWNYLTICQLCRPCDPVMGLEEIAPCTSKRKTQCRCQPGMFCAAWALECT
HCELLSDCPPGTEAELKGQRSLRGWM

40
>DEX0367_035.aa.3
KKREGGREKKGSGALIIVWVSISFLQGEGPNPVAGSWEPPKAHPYFPDLVQPLLPISGDVSPVSTGLPAA
PVLEAGVPQQQSPLDLTREPQLEPGEQSQVAHGTNGIHVTGGSMTITGNIYIYNGPVLGGPPGPGDLPAT
PEPPYPIPEEGDPGPPGLSTPHQEDGKAWHLAETEHCGATPSNRGPRNQFITHD

45
>DEX0367_036.aa.1
MATGLSEHHNMVWEVKTNQMPNAVQKLLLVMDKRASGMNDSLELLQCNENLPSSPGYNSCDEHMELDDLP
ELQAVQSDPTQSGMYQLSSDVSHQEYPRSSWNQNTSDIPETTYRENEVDWLTELANIATSPQSPLMQCSF
YNRSSPVHIIATSKSLHSYARPPPVSSSSKSEPAFPHHHWKEETPVRHERANSESESGIFCMSSLSDDDD
50   LGWCNSWPSTVWHCFLKAMTSHLWILLQFMC

>DEX0367_037.aa.1
MTGLALLYSGVFVAFWACALAVGVCYTIFDLGFRFDVAWFLTETSPFMWSNLGIGLAISLSVVGAAWGIY
ITGSSIIGGGVKAPRIKTKNLVSIIFCEAVAIYGIIMAIVISNMAEPFSATDPKAIGHRNYHAGYSMFGA
55   GLTVGLSNLFCGVCVGIVGSGAALADAQNPSLFVKILIVEIFGSAIGLFGVIVAILQVMNPLGKPLCPCP
QPSLTLLLEKLKCSPSLPITIDSPKQLPPPHFHLLVFSYRGSLFLSLIWCH

>DEX0367_038.aa.1
MAPARTMARARLAPAGIPAVALWLLCTLGLQGTQAGPPPAPPGLPAGADCLNSFTAGVPGFVLDTNASVS
60   NGATFLESPTVRRGWDCVRACCTTQNCNLALVELQPDRGEDAIAACFLINCLYEQNFVCKFAPREGFINY

LTREVYRSYRQLRTQGFGGSGIPKAWAGIDLKVQPQEPLVLKDVENTDWRLLRGDTDVRVERKDPNQVEL
WGLKEGTYLFQLTVTSSDHPEDTANVTVTVLSTKQTEDYCLASNKVGRCRGSFPRWYYDPTEQICKSFVY
GGCLGNKNNYLREEECILACRGVQGGPLRGSSGAQATFPQGPSMERRHPVCSGTCQPTQFRCSNGCCIDS
FLECDDTPNCPDASDEAACEKYTSGFDELQRIHFPSDKGHCVDLPDTGLCKESIPRWYYNPFSEHCARFT
5 YGGCYGNKNNFEEEQQCLESCRGISSEWASERVGMYGGRLSQWPPLCPQAFGSHPSILRAPGVGVGEDAS
VRSGALGSS

>DEX0367_039.aa.1
MVRILANGEIVQDDDPRVRTTTQPPRGSIPRQSFFNRGHGAPPGGPGPRQQQAGARLGAAQSPFNDLNRQ
10 LVNMGFPQWHLGNHAVEPVTSILLLFLLMMLGVRGLLLVGLVYLVSHLSQR

>DEX0367_039.aa.2
MVRILANGEIVQDDDPRVRTTTQPPRGSIPRQSFFNRGHGAPPGGPGPRQQQAGARLGAAQSPFNDLNRQ
LVNMGFPQWHLGNHAVEPVTSILLLFLLMMLGVRGLLLVGLVYLVSHLSQR

15

>DEX0367_039.aa.3
MIHSSLSVFTFQSFFNRGHGAPPGGPGPRQQQAGARLGAAQSPFNDLNRQLVNMGFPQWHLGNHAVEPVT
SILLLFLLMMLGVRGLLLVGLVYLVSHLSQR

20 >DEX0367_040.aa.1
MHARAAQCDGSVVAADPGDGTQLQNFTLDRSSVLVDGYSPNRNEPLTGNSDLPFWAVILIGLAGLLGLIT
CLICGVLVTTRRRKKEGEYNVQQQCPGYYQSHLDLEDLQ

>DEX0367_040.aa.2
25 MQHPGSRKFNTTERVLQGLLKPLFKSTSVGPLYSGCRLTLLRPEKHGAATGVDAICTLRLDPTGPGLDRE
RLYWELSQLTNSVTELGPYTLDRDSLYVNGFTHRSSVPTTSIPGTSAVHLETSGTPASLPGHTAPGPLLV
PFTLNFTITNLQYEEDMRHPGSRKFNTTERVLQGLLKPLFKSTSVGPLYSGCRLTLLRPEKRGAATGVDT
ICTHRLDPLNPGLDREQLYWELSKLTRGIIELGPYLLDRGSLYVNGFTHRNFVPITSTPGTSTVHLGTSE
TPSSLPRPIVPGPLLVPFTLNFTITNLQYEEAMRHPGSRKFNTTERVLQGLLRPLFKNTSIGPLYSSCRL
30 TLLRPEKDKAATRVDAICTHHPDPQSPGLNREQLYWELSQLTHGITELGPYTLDRDSLYVDGFTHWSPIP
TTSTPGTSIVNLGTSGIPPSLPETTATGPLLVPFTLNFTITNLQYEENMGHPGSRKFNITESVLQGLLKP
LFKSTSVGPLYSGCRLTLLRPEKDGVATRVDAICTHRPDPKIPGLDRQQLYWELSQLTHSITELGPYTLD
RDSLYVNGFTQRSSVPTTSTPGTFTVQPETSETPSSLPGPTATGPVLLPFTLNFTIINLQYEEDMHRPGS
RKFNTTERVLQGLLMPLFKNTSVSSLYSGCRLTLLRPEKDGAATRVDAVCTHRPDPKSPGLDRERLYWKL
35 SQLTHGITELGPYTLDRHSLYVNGFTHQSSMTTTRTPDTSTMHLATSRTPASLSGPTTASPLLVLFTINF
TITNLRYEENMHHPGSRKFNTTERVLQGLLRPVFKNTSVGPLYSGCRLTLLRPKKDGAATKVDAICTYRP
DPKSPGLDREQLYWELSQLTHSITELGPYTLDRDSLYVNGFTQRSSVPTTSIPGTPTVDLGTSGTPVSKP
GPSAASPLLVLFTLNFTITNLRYEENMQHPGSRKFNTTERVLQGLLRSLFKSTSVGPLYSGCRLTLLRPE
KDGTATGVDAICTHHPDPKSPRLDREQLYWELSQLTHNITELGPYALDNDSLFVNGFTHRSSVSTTSTPG
40 TPTVYLGASKTPASIFGPSAASHLLILFTLNFTITNLRYEENMWPGSRKFNTTERVLQGLLRPLFKNTSV
GPLYSGCRLTLLRPEKDGEATGVDAICTHRPDPTGPGLDREQLYLELSQLTHSITELGPYTLDRDSLYVN
GFTHRSSVPTTSTGVVSEEPFTLNFTINNLRYMADMGQPGSLKFNITDNVMKHLLSPLFQRSSLGARYTG
CRVIALRSVKNGAETRVDLLCTYLQPLSGPGLPIKQVFHELSQQTHGITRLGPYSLDKDSLYLNGYNEPG
PDEPPTTPKPATTFLPPLSEATTAMGYHLKTLTLNFTISNLQYSPDMGKGSATFNSTEGVLQHLLRPLFQ
45 KSSMGPFYLGCQLISLRPEKDGAATGVDTTCTYHPDPVGPGLDIQQLYWELSQLTHGVTQLGFYVLDRDS
LFINGHHTLQRQSTT

>DEX0367_040.aa.3
MRGRGRPGRCVLVDGYSPNRNEPLTGNSDLPFWAVILIGLAGLLGLITCLICGVLVTTRRRKKEGEYNVQ
50 QQCPGYYQSHLDLEDLQ

>DEX0367_040.aa.4
MHPPGFLSFLGYSPNRNEPLTGNSDLPFWAVILIGLAGLLGLITCLICGVLVTTRRRKKEGEYNVQQQCP
GYYQSHLDLEDLQ
55

>DEX0367_041.aa.1
MATHHTLWMGLALLGVLGDLQAAPEAQVSVQPNFQQDKFLGRWFSAGLASNSSWLREKKAALSMCKSVVA
PATDGGLNLTSTFLRKNQCETRTMLLQPAGSLGSYSYRSPHWGSTYSVSVVETDYDQYALLYSQGSKGPG
EDFRMATLYNKCMTEQ
60

>DEX0367_041.aa.2
MATHHTLWMGLALLGVLGDLQAAPEAQVSVQPNFQQDKFLGRWFSAGLASNSSWLREKKAALSMCKSVVA
PATDGGLNLTSTFLRKNQCETRTMLLQPAGSLGSYSYRSPREWGLHRPPGPSLGATLAGTTLGQPPAAEI
HGVGGDWGSTYSVSVVETDYDQYALLYSQGSKGPGEDFRMATLYSRTQTPRAELKEKFTAFCKAQGFTED

5   TIVFLPQTDKCMTEQ

>DEX0367_041.aa.3
MATHHTLWMGLALLGVLGDLQAAPEAQVSVQPNFQQDKFLGRWFSAGLASNSSWLREKKAALSMCKSVVA
PATDGGLNLTSTFLRKNQCETRTMLLQPAGSLGSYSYRSPHWGSTYSVSVVETDYDQYALLYSQAEPRPP

10   GLS

>DEX0367_041.aa.4
MATHHTLWMGLALLGVLGDLQAAPEAQVSVQPNFQQDKVRGFPASSPRATGPCQGKGTFRLGLPPGRSER
SPAVPGSAGQGLSGRAGRRLGSRPRRLPARALPGHRVPSPLMGHADTGPHTRPRQPDTSTPVGTRPPEDT

15   RAHVPHLGARTRAGGAQGWRQTLRARWGLGGTRTAQTAGDARSRPGAARGSAGARVPAPRAPPWRRGEPQ
RSAELSRRPAPIPARNATSSAARCMGQALSQGTESGPGAEGPKLAGGRRARFLGRWFSAGLASNSSWLRE
KKAALSMCKSVVAPATDGGLNLTSTFLRKNQCETRTMLLQPAGSLGSYSYRSPHWGSTYSVSVVETDYDQ
YALLYSQGSKGPGEDFRMATLYSRTQTPRAELKEKFTAFCKAQGFTEDTIVFLPQTDKCMTEQ

20   >DEX0367_041.aa.5
MATHHTLWMGLALLGVLGDLQAAPEAQVSVQPNFQQDKVRGFPASSPRATGPCQGKGTFRLGLPPGRSER
SPAVPGSAGQGLSGRAGRRLGSRPRRLPARSPPWAPRPVSPDGPRRHRATHAPTPARHVHPCGHATPRGH
TSARSTPGCQDTGGWGTGMATNTPCAVGVGRDAHRTDSRRRALSPGSCSGKRRSAGPRAARPSLASRRTP
AVRRAEPKTRPDPRQECDVLCRPLYGPGAQPGHRIGTGGGGAEAGWWAACEGEGLSSGGAWPDGGAGCQG

25   RGQLLGRRCEGRGHLLGRGLRGGGQFLGRGVRGVAARGIGRGGGAGLETGGVDGRGAPAGRRRWVRRVLA
DAGGGRSPQFPGALVQRGPRLQLELAPGEEGGVVHVQVCGGPCHGWWPQPDLHLPQEKPV

>DEX0367_041.aa.6
MAAGPMAAEPCGPHALVALAGLVTGIPTHHPRVYNIHSRTVTRYPANSIVVVGGCPVCRVGVLEDCFTFL

30   GIFLAIILFRIGPAAIGQWQPPNGSRTQTPRAELKEKFTAFCKAQGFTEDTIVFLPQTDKCMTEQ

>DEX0367_041.aa.7
ALLEAWARDRGVSVQVRTSLPQPLHEEPPPWGTWRPGAHSVPGPSSSQDVGLQPGGGHRVEGAHGGYRGT
NHTGLRHSLLGVDSLLLAEVEKDPLFVSSAQGEVGGDGGSVQFGGSVKTSSALREEQEAQWENWPKSGVL

35   TTAPGFFLGR

>DEX0367_042.aa.1
MDIYDTQTLGVVVFGGFMVVSAIGIFLVSTFSMKETSYEEALANQRKEMAKTHHQKVEKKKKEKTVEKKG
KTKKKEEKPNGKIPDHDPAPNVTVLLREPVRAPAVAVAPTPVQPPIIVAPVATVPAMPQEKLASSPKDKK

40   KKEKKVAKVEPAVSSVVNSIQVLTSKAAILETAPKEVPMVVVPPVGAKGNTPATGTTQGKKAEGTQNQSK
KAEGAPNQGRKAEGTPNQGKKTEGTPNQGKKAEGTPNQGKKAEGTPNQGKKAEGAQNQGKKVDTTPNQGK
KVEGAPTQGRKAEGAQNQAKKVEGAQNQGKKAEGAQNQGKKGEGAQNQGKKAEGAQNQGKKAEGAQNQGK
KAEGAQNQGQKGEGAQNQGKKTEGAQGKKAERSPNQGKKGEGAPIQGKKADSVANQGTKVEGITNQGKKA
EGSPSEGKKAEGSPNQGKKADAAANQGKKTESASVQGRNTDVAQSPEAPKQEAPAKKKSGSKKKGEPGPP

45   DADGPLYLPYKTLVSTVGSMVFNEGEAQRLIEILSEKAGIIQDTWHKATQKGDPVAILKRQLEEKEKLLA
TEQEDAAVAKSKLRELNKEMAAEKAKAAAGEAKVKKQLVAREQEITAVQARMQASYREHVKEVQQLQGKI
RTLQEQLENGPNTQLARLQQENSILRDALNQATSQVESKQNAELAKLRQELSKVSKELVEKSEAVRQDEQ
QRKALEAKAAAFEKQVLQLQASHRESEEALQKRLDEVSRELCHTQSSHASLRADAEKAQEQQQQMAELHS
KLQSSEAEVRSKCEELSGLHGQLQEARAENSQLTERIRSIEALLEAGQARDAQDVQASQAEADQQQTRLK

50   ELESQVSGLEKEAIELREAVEQQKVKNNDLREKNWKAMEALATAEQACKEKLHSLTQAKEESEKQLCLIE
AQTMEALLALLPELSVLAQQNYTEWLQDLKEKGPTLLKHPPAPAEPSSDLASKLREAEETQSTLQAECDQ
YRSILAETEGMLRDLQKSVEEEEQVWRAKVGAAEEELQKVYAALPASRRRGAHEAQGRVDRIVSVTCETR
VSRWGFSTTPYEAQLRDDSSVRPPGGPRLGRHCQMAPG

55   >DEX0367_042.aa.3
MDIYDTQTLGVVVFGGFMVVSAIGIFLVSTFSMKETSYEEALANQRKEMAKTHHQKVEKKKKEKTVEKKG
KTKKKEEKPNGKIPDHDPAPNVTVLLREPVRAPAVAVAPTPVQPPIIVAPVATVPAMPQEKLASSPKDKK
KKEKKVAKVEPAVSSVVNSIQVLTSKAAILETAPKEVPMVVVPPVGAKGNTPATGTTQGKKAEGTQNQSK
KAEGAPNQGRKAEGTPNQGKKTEGTPNQGKKAEGTPNQGKKAEGTPNQGKKAEGAQNQGKKVDTTPNQGK

60   KVEGAPTQGRKAEGAQNQAKKVEGAQNQGKKAEGAQNQGKKGEGAQNQGKKAEGAQNQPPDVTVLLREPV

RAPAVAVAPTPVQPPIIVAPVATVPAMPQEKLASSPKDKKKKEKNVAKVEPAVSSVVNSIQSSHFEGCQV
GSMVFNEGEAQRLIEILSEKAGIIQDTWHKATQKGDPVAILKRQLEEKEKLLATEQEDAAVAKSKLRELN
KEMAAEKAKAAAGEAKVKKQLVAREQEITAR

5 >DEX0367_043.aa.1
GRSQRSSPCSAPLQGPGALGLRTQLLLPPWSSTWEQVSSWGVWTGGAGGRTQAQKLPAPTTQLLSTALEP
TSQKPGVGAGHGGDPKLSPHKVQGRSEAGAGPGPKQGHHSSSDSSSSSSDSDTDVKSHAAGSKQHESIPG
KAKKPKVKKKEKGKKEKGKKKEAPH

10 >DEX0367_043.aa.2
GRSQRSSPCSAPLQGPGALGLRTQLLLPPWSSTWEQVSSWGVWTGGAGGRTQAQKLPAPTTQLLSTALEP
TSQKPGVGAGHGGDPKLSPHKVQGRSEAGAGPGPKVSRLITGCGGAGRLGLPLTPGSCLRPPTSGGWVRG
AASLPHLHPARTPQLFRLQQQLQRFGHGCEGKGLSPASPSTCPAPQRGVPALGLAGRVRGVVPLSRFQPR
TISSLPSLPSAPQSHAAGSKQHESIPGKAKKPKVKKKEKGKKEKGKKKEAPH
15

>DEX0367_044.aa.1
MAQLCGLRRSRAFLALLGSLLLSGVLAADRERSIHDFCLVSKVVGRCRASMPRWWYNVTDGSCQLFVYGG
CDGNSNNYLTKEECLKKCATVTENATGDLATSRNAADSSVPSAPRRQDSEDHSSDMFNYEEYCTANAVTG
PCRASFPRWYFDVERNSCNNFIYGGCRGNKNSYRSEEACMLRCFQRELPWPWAKGGRGAARGGTPRGAQG
20 TEP

>DEX0367_044.aa.2
MAQLCGLRRSRAFLALLGSLLLSGVLAADRERSIHDFCLVSKVVGRCRASMPRWWYNVTDGSCQLFVYGG
CDGNSNNYLTKEECLKKCATVTENATGDLATSRNAADSSVPSAPRRQDSEDHSSDMFNYEEYCTANAVTG
25 PCRASFPRWYFDVERNSCNNFIYGGCRGNKNSYRSEEACMLRCFRQQENPPLPLGSKGKWPLTLLLPSAC
LLPSLTELSPAQLWFTLSFTVNIILAESHVSA

>DEX0367_045.aa.1
MDRPPGQVPGHLGQCDVSGWQSDAGPAGSQENSTLVPEERWKF
30

>DEX0367_045.aa.2
MDRSRPRYLAILGSVTFLAGNRMLAQQAVKRTAH

>DEX0367_046.aa.1
35 MESEELNGXXKAIPVXXDLNAPSDWDSRGKDSYETRXLDXKSAEXKXKKKSRLYKRKANDESKEQADVIQ
RKELAKDSREANSHEFXSKEDMLVVDAKSKEEEKHLKFRISHELDSASSEFN

>DEX0367_047.aa.1
KSGINAEGPLRPGAAILGLLGLASVGDSRT .
40

>DEX0367_048.aa.1
MHQGREHSAQTIDSDYLLTVCRRHAGGLKRGCILLPSLAGYGISSFPFIGYEVYSSIYQNL

>DEX0367_049.aa.1
45 SGGFTLRSTGSAAGRGLRKLRPTLSRGACDVGRSVWKLRPSHCRVGRGARGIFHAPPLCRLKCVGGKRWA
GCHLAPPFSCGVGQAAASSSCSSHRLHSDPSPAAWILLPRRGSFSRNLRARPQSVPAASRRTRWLQASLT
PGFLAKRETEGERGEDERTGNGR

>DEX0367_050.aa.1
50 MACPFELLFSKPPWGPSPPLLPKFWSPQKTQGNIVWSAQQIKGNAQTPMGAPTLQPVRARQRTPRPWGTW
GSLQSCQKKPCAKSGGLPRGSVENIVPYGDMEGGQCRGEATKATINGDSRRSSQQGAHGYEITAEKPPVG
RNRVSNTLGCPQGTGTGAEHGAVLGVTGVLRKKYPGLVLEVPRDKLTAAGMKTDEGPH

>DEX0367_051.aa.1
55 MVPEPKIIDAALRACRRLNDFASTVRILEVVKDKAGPHKEIYPYVIQELRPTLNELGISTPEELGLDKV

>DEX0367_052.aa.1
MLPPATAQLCLFDPRAVAGLYISLYDHKLPANESICVDNLTYNILSKIMHFRQTDISLFSIYSFF

60 >DEX0367_052.aa.2

MKLLTPPGADSHLNSTPSKLRRSPAKN

>DEX0367_053.aa.1
MAAMASLGALALLLLSSLSRCSAEACLEPQITPSYYTTSDAVISTETVFIVEISLTCKNRVQNMALYADV
5   GGKQFPVTRGQDVGRYQVSWSLDHKSAHAGTYEVRFFDEESYSLLRKAQRNNEDISIIPPLFTVSVDHRV
SGLVPPPFWGCWAE

>DEX0367_053.aa.2
ATYASLFLVCPSCSWELVFFSSRQRRGDGGDGISRRPGAAPAVQPLPLLRKRWGQSSPSVNWPGAETLSL
10   ILKAPVHCLPAAPTPDVSAEDAFSNSSRPFVALNVRLAWPPGLRVIGPSVAEACLEPQITPSYYTTSDAV
ISTETVFIVEISLTCKNRVQNMALYADVGGKQFPVTRGQDVGRYQVSWSLDHKSAHAGTYEVRFFDEESY
SLLRKAQRNNEDISIIPPLFTVSVDHRGTWNGPWVSTEVLAAAIGLVIYYLAFSAKSHIQA

>DEX0367_054.aa.1
15   MKSVIFQFGIKVSFSVAAKCLVMKAEMNGSKLGRRAKPEGALQNNDGLYDPDCDESGLFKAKQCNGTSMC
WCVNTAGVRRTDKDTEITCSERVRTYWIIIELKHKAREKPYDSKSLRTALQKEITTRYQLDPKFITSILY
ENNVITIDLVQNSSQKTQNDVDIADVAYYFEKDVKGESLFHSKKMDLTVNGEQLDLDPGQTLIYYVDEKA
PEFSMQGLKAGVIAVIVVVVIAVVAGIVVLVISRKKRMAKYEKAEIKEMGEMHRELNA

20   >DEX0367_055.aa.1
ALSSGRPGRYSVWIGGSILASLSTFQQMWISKQEYDESGPSIVHRKCF